

Katrin Hillgruber

**Persönlichkeitsbeschreibende Verfahren in der Personaldiagnostik –
eine korrelative Vergleichsstudie zwischen MBTI, EPQ und Neo-FFI**

eingereicht als

DIPLOMARBEIT

an der

Hochschule Mittweida

University of Applied Sciences

Fachbereich Wirtschaftswissenschaften

Mittweida, 2010

Erstprüfer: Prof. Dr. rer. pol. Ulla Meister
Zweitprüfer: Dipl. Psych. Frank Schumann

Katrin Hillgruber

**Persönlichkeitsbeschreibende Verfahren in der Personaldiagnostik –
eine korrelative Vergleichsstudie zwischen MBTI, EPQ und Neo-FFI**

eingereicht als

DIPLOMARBEIT

an der

Hochschule Mittweida
University of Applied Sciences

Fachbereich Wirtschaftswissenschaften

Mittweida, 2010

Erstprüfer: Prof. Dr. rer. pol. Ulla Meister

Zweitprüfer: Dipl. Psych. Frank Schumann

Bibliographische Beschreibung:

Hillgruber, Katrin:

Persönlichkeitsbeschreibende Verfahren in der Personaldiagnostik - eine korrelative Vergleichsstudie zwischen MBTI, EPQ und Neo-FFI.

- 2010 – 119 Seiten,

Mittweida, Hochschule Mittweida (FH), Fachbereich Wirtschaftswissenschaften

Diplomarbeit, 2010

Inhaltsverzeichnis

Abbildungsverzeichnis.....	IV
Tabellenverzeichnis.....	V
Abkürzungsverzeichnis.....	VI
1 Einleitung.....	1
2 Persönlichkeitsverfahren in der Berufseignungsdiagnostik.....	3
2.1 Historischer Abriss der Eignungsdiagnostik	3
2.2 Theoretische Grundlagen	5
2.2.1 Das Eigenschaftskonzept.....	5
2.2.2 Stabilität und Veränderung von Persönlichkeitsmerkmalen	7
2.2.3 Einordnung von Persönlichkeitsverfahren	9
2.3 Notwendigkeit von Persönlichkeitsverfahren in der psychologischen Berufseignungsdiagnostik.....	10
2.4 Personalpolitische Ausgangslage	15
2.5 Der Einsatz von Persönlichkeitsverfahren im internationalen Vergleich	17
2.6 Anforderungen an Testverfahren	20
2.6.1 Übersicht.....	20
2.6.2 Hauptgütekriterien.....	20
2.6.3 Der Zusammenhang von Objektivität, Reliabilität und Validität.....	26
2.6.4 Nebengütekriterien.....	27
2.7 Anforderungen an Unternehmen	30
2.7.1 Juristische Aspekte der Berufseignungsdiagnostik.....	30
2.7.2 Ethische Aspekte der Berufseignungsdiagnostik	35
2.7.3 Durchführung testgestützter Eignungsuntersuchungen	36
2.8 Motivationale Einflüsse und Fehlerfaktoren bei der Bearbeitung von Persönlichkeitsverfahren	38
2.8.1 Beantwortungen im Sinne sozialer Erwünschtheit.....	39
2.8.2 Zustimmungstendenzen	41
2.8.3 Bevorzugung extremer oder mittlerer Antwortkategorien.....	41

3	Theoretische Darstellung der ausgewählten Testverfahren	43
3.1	Das NEO-Fünf-Faktoren Inventar (NEO-FFI) nach Costa und McCrae.....	43
3.1.1	Grundkonzept	43
3.1.2	Theoretische Begründung des Verfahrens	44
3.1.3	Die Entwicklung des Verfahrens.....	45
3.1.4	Skalen und Interpretationshinweise.....	45
3.1.5	Normen und Gütekriterien	46
3.2	Der Myers-Briggs Typenindikator (MBTI).....	48
3.2.1	Grundkonzept	48
3.2.2	Theoretischer Hintergrund.....	49
3.2.3	Entwicklung des Verfahrens.....	52
3.2.4	Skalen und Interpretationshinweise.....	53
3.2.5	Normen und Gütekriterien	54
3.3	Der Eysenck Personality Questionnaire.....	56
3.3.1	Grundkonzept	56
3.3.2	Theoretische Begründung des Verfahrens	57
3.3.3	Die Entwicklung des Verfahrens.....	59
3.3.4	Skalen und Interpretationshinweise.....	61
3.3.5	Normen und Gütekriterien	62
4	Empirische Belege und Fragestellungen	64
4.1	Empirische Belege.....	64
4.1.1	Die Neuinterpretation des Myers-Briggs Typenindikators aus der Perspektive des Fünf-Faktoren- Modells der Persönlichkeit	64
4.1.2	Der Zusammenhang zwischen dem revidierten NEO- Persönlichkeitsinventar und dem Myers-Briggs Typenindikator.....	65
4.1.3	Der Zusammenhang zwischen dem Myers-Briggs Typenindikator und dem Eysenck Personality Questionnaire unter erwachsenen Kirchgängern.....	66
4.1.4	Die Untersuchung zwischen drei Verfahren der Selbsteinschätzung bezüglich Persönlichkeit, Typen und Störungen.....	68
4.2	Zentrale Fragestellungen.....	68

5	Methodisches Vorgehen	70
5.1	Stichprobenbeschreibung	70
5.2	Datenerhebung	70
5.3	Datenaufbereitung	71
5.4	Datenauswertung	72
5.4.1	Test auf Normalverteilung	72
5.4.2	Korrelationsanalysen	73
6	Ergebnisdarstellung und Interpretation	76
6.1	Test auf Normalverteilung	76
6.2	Interkorrelationen in den einzelnen Verfahren	78
6.3	Korrelativer Vergleich zwischen den Verfahren	83
6.4	Zusammenfassende Ergebnisse	92
6.5	Methodische Anmerkungen und Einschränkungen der vorliegenden Untersuchung	94
7	Perspektiven	95
	Anhang	VII
	Literaturverzeichnis	XIV
	Eigenständigkeitserklärung	XXI

Abbildungsverzeichnis

Abb. 1: Die drei Quellen diagnostischer Information	9
Abb. 2: Determinanten des Berufserfolges	11
Abb. 3: Woran scheitern High Potentials im Unternehmen?	12
Abb. 4: Einsatz von Tests oder Fragebögen bei der Personalauswahl	19
Abb. 5: Die klassischen testtheoretischen Gütekriterien	20
Abb. 6: Interdependenz der Gütekriterien	26
Abb. 7: Das BSG-Modell.....	29
Abb. 8: Schichtenmodell der Persönlichkeit im Rahmen des allgemeinen Persönlichkeitsrechts	33
Abb. 9: Eysencks Persönlichkeitszirkel.....	59

Tabellenverzeichnis

Tab. 1: Anwendung verschiedener Auswahlverfahren für Führungskräfte im europäischen Vergleich	18
Tab. 2: Die fünf Merkmalsbereiche des NEO-FFI	45
Tab. 3: Die Reliabilitäten des NEO-FFI	47
Tab. 4: Die 16 Typen des MBTI	52
Tab. 5: Die Skalen des EPQ-RK	61
Tab. 6: Prüfung der Daten des NEO-FFI auf Normalverteilung	76
Tab. 7: Prüfung der Daten des EPQ-RK auf Normalverteilung	76
Tab. 8: Prüfung der Daten des MBTI auf Normalverteilung	77
Tab. 9: Die Interkorrelationen im NEO-FFI	79
Tab. 10: Die Interkorrelationen im EPQ-RK	80
Tab. 11: Verteilung der Präferenzen des MBTI	81
Tab. 12: Die Interkorrelationen im MBTI	82
Tab. 13: Die Korrelationen zwischen dem EPQ-RK und dem MBTI	84
Tab. 14: Die Korrelationen zwischen dem NEO-FFI und dem MBTI	87
Tab. 13: Die Korrelationen des EPQ-RK und des NEO-FFI	91

Abkürzungsverzeichnis

Abb.	Abbildung
bzw.	beziehungsweise
d.h.	das heißt
EPQ-RK	Eysenck Personality Questionnaire – revidiert/ Kurzform
etc.	et cetera
f.	folgende
ff.	fortfolgende
gem.	gemäß
MBTI	Myers-Briggs-Typen-Indikator
NEO-FFI	NEO-Fünf-Faktoren-Inventar
Nr.	Nummer
o.ä.	oder ähnliches
Rn.	Randnummer
S.	Seite
sog.	sogenannt
Tab.	Tabelle
u.a.	unter anderem
URL	Uniform Ressource Locator
usw.	und so weiter
v.a.	vor allem
v. Chr.	vor Christus
vgl.	vergleiche
z.B.	zum Beispiel
zit. n.	zitiert nach

1 Einleitung

„Young man, we hire people because of their knowledge and experience, but we fire them because of their personality.“¹

Viele Millionen Auswahlentscheidungen werden jährlich getroffen. Berufsbezogene Entscheidungen gehören dabei sowohl für die betroffenen Personen als auch für die Organisationen zu den wichtigsten Selektionshandlungen überhaupt. Die Qualität dieser Entscheidungen ist dabei so vielfältig wie die zur Entscheidungsfindung verwendeten Methoden. Berufliche Entscheidungen sind nicht nur am Anfang des Berufslebens und beim Eintritt in ein Unternehmen oder eine Organisation zu treffen. Vielmehr stehen in den meisten beruflichen Entscheidungsverläufen in bestimmten zeitlichen Abständen neue Entscheidungen über die passendste Orientierung unter den verschiedenen Möglichkeiten an.²

Beruflicher Erfolg hängt von vielen Faktoren ab. Der familiäre und soziale Hintergrund des Menschen prägt seine Einstellungen und Erwartungen; die Ausbildung schafft die Grundlage für die weiteren Entwicklungsmöglichkeiten. Fachliche Qualifikationen sind also zweifelsohne eine wichtige Voraussetzung für den langfristigen beruflichen Erfolg. Doch spielt auch die Persönlichkeit eines Menschen in ihren verschiedenen Facetten eine gleichermaßen bedeutende Rolle, wie schon ein erfahrener Personalleiter eines großen englischen Konzerns im obigen Zitat feststellte.

Die psychologische Berufseignungsdiagnostik ist bestrebt, Zusammenhänge zwischen menschlichen Merkmalen und beruflichen Erfolg zu entdecken und daraus Methoden zu entwickeln, um beides zu messen und zueinander in Beziehung zu bringen. Die sachgemäße Anwendung dieser Methode lässt nachweislich bessere Prognosen des Berufserfolges zu als jede andere Vorgehensweise. Damit stellen die unterschiedlichen Verfahren der Berufseignungsdiagnostik – auch wenn es Verbesserungsbedarf bei den einzelnen Verfahren gibt - die angemessenste derzeit bekannte Grundlage bei der Auswahl und Entwicklung von Mitarbeitern und der Berufsberatung dar.³

Ziel der vorliegenden Arbeit ist es zu belegen, dass die Persönlichkeit eines Menschen einen wichtigen beeinflussenden Faktor des beruflichen Werdeganges darstellt und im

¹ vgl. Nußbaum/ Neumann, 1995, S.127

² vgl. Schuler, 2000, S. 5 f.

³ vgl. ebenda

Hinblick auf ganzheitliches Personalmanagement nicht vernachlässigt werden darf. Es soll darauf eingegangen werden, welchen Anforderungen ein Unternehmen gerecht werden muss, um seriös und erfolgreich mit Persönlichkeitsverfahren im Personalmanagement arbeiten zu können.

Im empirischen Teil der Arbeit werden drei persönlichkeitsbeschreibende Verfahren vorgestellt. Um mögliche Zusammenhänge zwischen den Verfahren, denen unterschiedliche theoretische Modelle zugrunde liegen, aufzudecken, werden diese einem korrelativem Vergleich unterzogen.

Abschließend sollen die Ergebnisse der Arbeit zusammengefasst und mögliche Perspektiven aufgezeigt werden.

Anmerkung:

Aus Gründen der besseren Lesbarkeit wird in dieser Arbeit darauf verzichtet, Begriffe, welche Personen oder Gruppen betreffen, in männlicher und weiblicher Form aufzuführen. Falls es keine expliziten Hinweise gibt, sind grundsätzlich beide Geschlechter gemeint.

2 Persönlichkeitsverfahren in der Berufseignungsdiagnostik

2.1 Historischer Abriss der Eignungsdiagnostik

Die ersten Ansätze des Selektionsgedankens hinsichtlich Mitarbeiterereinstellungen tauchten schon vor 4000 Jahren in China auf.⁴ Etwa 3000 Jahre später, während der Ming-Dynastie (1386-1644 v. Chr.), wurden schon Vorformen von Assessment Centern zur Auswahl von Führungskräften durchgeführt.⁵ Dieses Auswahlverfahren der Chinesen wurde Anfang des 18. Jahrhunderts von britischen Missionaren und Diplomaten teilweise übernommen. Es bildete später den Ausgangspunkt für die Auswahl britischer Staatsbeamter.⁶

Die beschriebenen Beispiele beruhten jedoch auf keinen mess- und testtheoretischen Erkenntnissen. Ende des 19. Jahrhunderts wurden erstmals systematische Vorgehensweisen angewandt.⁷ Zu dieser Zeit beschäftigte sich der britische Forscher Sir Francis Galton mit dem Zusammenhang von z.B. physischen Merkmalen und der Erledigung bestimmter Aufgaben. James McKeen Cattell begann 1890 in Amerika, teilweise auf den Studien von Galton aufbauend, mit der Messung motorischer und sensorischer Fähigkeiten mithilfe von Testbatterien und der Untersuchung menschlichen Denkvermögens. Weiterhin forschten Hermann Ebbinghaus, Alfred Binet und William Stern im Grundlagenbereich der Intelligenz.

Schließlich zeigte Hugo Münsterberg (1912), gründend auf diesen Forschungen und Ergebnissen, dass zwischen im Labor erhobenen Daten und Leistungen am Arbeitsplatz tatsächlich Zusammenhänge bestehen. Damit legte er den Grundstein der Eignungsdiagnostik.⁸

Die wissenschaftlich-psychologisch fundierte und nach psychometrischen Maximen konstruierte Eignungsdiagnostik hielt zu Anfang des 20. Jahrhunderts in Deutschland Einzug und wurde primär von staatlichen Institutionen und großen Unternehmen (z.B. Siemens, AEG) genutzt. Dabei lagen die Schwerpunkte auf Papier- und Bleistifttests sowie Arbeitsproben. In den Vereinigten Staaten wurden zu dieser Zeit, hauptsächlich

⁴ vgl. Wiggins, 1973, S. 515

⁵ vgl. http://en.wikibooks.org/wiki/Applied_History_of_Psychology/Models_of_Assessment, verfügbar am 15.05.2010

⁶ vgl. Wiggins, 1973, S.516

⁷ vgl. Schuler, 2000, S. 15

⁸ vgl. DuBois, 1970, S. 3 f.

im militärischen Bereich, bereits biographische Fragebögen, Persönlichkeits- und Intelligenztests eingesetzt.⁹ Ende der zwanziger Jahre in der Weimarer Republik fanden bei der Offiziersauswahl multiple Auswahlverfahren ihre Anwendung, die die umfassende Persönlichkeitserfassung zum Ziel hatten.¹⁰ Diese gelten als Vorläufer der heutigen Assessment Center zur Führungskräfteauswahl.

Basierend auf diesen deutschen Auswahlverfahren wurde während des Zweiten Weltkrieges die Auswahl britischer Offiziere neu konzipiert und später bei der Auswahl britischer Staatsbeamter ebenfalls eingesetzt. Der Eintritt der Vereinigten Staaten in den Zweiten Weltkrieg führte zu einem gestiegenen Bedarf an Geheimdienstmitarbeitern. Zur Selektion wurde nun auch in Übersee das ehemals deutsche, von den Briten adaptierte, Auswahlverfahren genutzt. Getestet worden zudem annähernd alle 14 Millionen amerikanische Kriegsteilnehmer. Der Einsatz psychologischer Testverfahren, auch „Herrespsychotechnik“ genannt, fand dann vor allem zur Auswahl von Kraftfahrern, Piloten, Funkern u.a. statt.¹¹ Hier erreichte das Testen einen ersten Höhepunkt. Auch im industriellen Bereich wurden zunehmend Testungen und andere systematische Verfahren zur Bewerberauswahl angewandt. Um bei den Auswahlverfahren ein gewisses Maß an Qualität zu sichern, wurde in den Vereinigten Staaten von 1954 bis 1966 durch die American Psychological Association Standards für die Entwicklung und Evaluation von Tests festgeschrieben. In Deutschland vernachlässigte man zu jener Zeit die Gütekriterien und setzte auf die Anwendung explorativer Gesprächstechniken.¹²

In den Sechzigern und Anfang der Siebziger spricht man von einer internationalen Krise der Berufseignungsdiagnostik. Durch gesellschaftspolitische Einflüsse wurde die Test-Fairness kritisiert und diskutiert. Außerdem erkannten die Forscher die Komplexität der Umwelt und stellten die einfachen Auswahlverfahren in Frage.¹³ Die ethisch-politische Diskussion dauerte in Deutschland, begünstigt durch die Einführung von Hochschulzulassungstests, bis in die Achtziger Jahre an. Seit dieser Zeit gewinnen eignungsdiagnostische Verfahren bei der Mitarbeiterauswahl in der Wirtschaft immer mehr an Bedeutung.¹⁴

⁹ vgl. Schuler, 2000, S. 15

¹⁰ vgl. Wiggins, 1973, S. 516

¹¹ vgl. Obermann, 2009, S. 20 f..

¹² vgl. Hartmann/ Haubl/ Neuberger, 1984; S. 79

¹³ vgl. http://en.wikibooks.org/wiki/Applied_History_of_Psychology/Models_of_Assessment,
verfügbar am 16.05.2010

¹⁴ vgl. Hossiep, 1995, S.1

Mit der Entwicklung der Informationstechnologie tritt auch die computergestützte Eignungsdiagnostik immer mehr in den Vordergrund. Computer werden dabei sowohl bei der Konstruktion, der Erfassung, der Auswertung als auch bei der Entscheidungsfindung eingesetzt. Auf dem ersten deutschsprachigen Kongress für Berufseignungsdiagnostik 1991 ergibt die Bestandsaufnahme, dass computerunterstützte Verfahren, Simulationen und Diagnosesysteme bereits die größere Gruppe unter den vorhandenen Verfahrenstypen bilden.¹⁵

Die Gesellschaft und die Arbeitswelt befinden sich seit einigen Jahren in einem tiefgreifendem Wandel. Man wendet sich vom reinen Produktionsdenken der vergangenen Jahre ab und geht hin zum Innovationsdenken und einer Wissensgesellschaft. Gerade europäische Unternehmen müssen sich Wettbewerbsvorteile sichern. Dahin gehend gewinnt das Humankapital eines Unternehmens und seine Auswahl eine immer größere Bedeutung.¹⁶

Die Eignungsdiagnostik gewinnt daher stetig an Bedeutung in der Personalarbeit, nicht nur in Bezug auf die Personalbeschaffung, sondern auch in Bezug auf die Personalentwicklung. Die Personalentscheidungen in den Unternehmen werden vermehrt durch Psychologen unterstützt.

2.2 Theoretische Grundlagen

2.2.1 Das Eigenschaftskonzept

Die Begriffe „Eigenschaft“ bzw. „trait“ werden in der Persönlichkeitspsychologie nicht einheitlich definiert – schon 1937 führte Allport fünfzig verschiedene Definitionen auf. Es existiert keine allgemein anerkannte Definition. Statt einer verbindlichen Begriffsbestimmung fasst man einzelne Elemente verschiedener Definitionen zusammen.¹⁷ Amelang und Bartussek vermuten, dass die Mehrzahl der empirisch tätigen Persönlichkeitsforscher einer Arbeitsgrundlage zustimmen könnte, in der Eigenschaften als „relativ breite und zeitlich stabile Dispositionen zu bestimmten

¹⁵ vgl. Schuler, 2000, S. 18

¹⁶ vgl. Etzel/ Küppers, 2002, S.11

¹⁷ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 23

Verhaltensweisen“ verstanden werden, „die konsistent (also durchweg) in verschiedenen Situationen auftreten“.¹⁸

Eine Eigenschaft kann man in der Regel nicht aus einer singulären Verhaltensweise ableiten. Erst das systematische gemeinsame Auftreten bestimmter Handlungen macht eine Beschreibung dieser Verhaltenstendenzen durch einen Eigenschaftsbegriff sinnvoll. Zum Beispiel kann das sofortige Aufräumen des Schreibtisches nach getaner Arbeit, die Pflege effektiver Ablagesysteme und das pünktliche Einhalten aller Termine auf eine „hoch ausgeprägte Gewissenhaftigkeit“ hindeuten. Zu beachten ist jedoch, dass das Eigenschaftskonzept eher als probabilistisches denn als kausales Modell zu verstehen ist, denn mit jeder Zusammenfassung und Abstraktion im Sinne einer ökonomischen wissenschaftlichen Modellbildung, geht auch ein Informationsverlust einher. Das freundliche Grüßen von Bekannten ist weder eine hinreichende noch eine notwendige Bedingung für das Vorliegen einer hohen Ausprägung von „Freundlichkeit“.¹⁹

Der klassische eigenschaftstheoretische Ansatz, der vor allem mit den Namen Guilford, Cattell und Eysenck verknüpft ist, wurde 1968 von Mischel heftig angegriffen. Hauptkritikpunkt stellte der Widerspruch dar, dass Eigenschaftstheorien die situationsübergreifende Konsistenz (völlige Gleichförmigkeit) des Verhaltens voraussetzen, durch die Forschung aber kaum Belege dafür gefunden werden konnten. Gegen die traditionelle Eigenschaftsforschung sprechen weiterhin die geringen Validitäten von Persönlichkeitsverfahren, die in der Höhe der Koeffizienten nur selten den Wert von .30 übertreffen. Ein solcher Wert ist bestenfalls zur Beschreibung von Gruppen geeignet, aber nicht zur Vorhersage individuellen Verhaltens. Seitdem wurden zahlreiche Bemühungen zur Entkräftung der Argumentation Mischels unternommen, in denen belegt wurde, dass die Argumente Mischels dem Eigenschaftskonzept nicht wirklich entgegenstehen. Schließlich haben sich Eigenschaftstheoretiker und die sogenannten Situationisten immer mehr angenähert, was das Zusammenwirken von Person und Situation betrifft. Die Eigenschaftstheoretiker haben akzeptiert, dass die Situation eine bedeutsame Verhaltensdeterminante ist, die Situationisten dagegen haben in einer Reihe von Bereichen Persönlichkeitseigenschaften als angemessene Kategoriensysteme für Verhaltenstendenzen aufgenommen.²⁰

¹⁸ Amelang/ Bartussek, 2006, S. 45 f.

¹⁹ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 23 f.

²⁰ vgl. ebenda

Aus der langjährigen Debatte um die Konsistenz menschlichen Verhaltens lassen sich verschiedene Konsequenzen für den Einsatz von Persönlichkeitsverfahren bei Personalentscheidungen ableiten: Persönlichkeitseigenschaften können durchaus eine gewisse Prädiktionskraft für zukünftiges Verhalten haben, denn menschliches Verhalten ist nicht, wie Mischel 1968 postuliert hat, so stark durch die Situation determiniert. Dass Persönlichkeitseigenschaften keine Prädiktionskraft für einzelne Verhaltensakte haben ist für die Berufseignungsdiagnostik sowieso nur von geringem Belang. Bei der Frage nach der beruflichen Bewährung steht besonders das typische zukünftige Verhalten im Zentrum der Aufmerksamkeit.²¹

2.2.2 Stabilität und Veränderung von Persönlichkeitsmerkmalen

„In most of us, by the age of thirty, the character has set like plaster, and will never soften again.“²²

Diese noch heute vielzitierte Bemerkung macht William James schon 1890 in seinem Buch „Principles of Psychology“. Die Diskussion um Stabilität und Veränderung von Persönlichkeit über die Zeit wird noch immer engagiert geführt.

Im Allgemeinen sind Persönlichkeitstraits so konzeptualisiert und definiert, dass sie explizit als zeitlich überdauernde und stabile Verhaltensdispositionen verstanden werden. Paul Costa und Robert McCrae – zwei US-amerikanische Forscher – haben sich um den Nachweis einer langfristigen Stabilität von Persönlichkeitseigenschaften bemüht. Sie haben mit Hilfe von Längsschnittstudien vor allem Belege für die zeitliche Stabilität der Faktoren des Big-Five-Modells erbracht, das zu einem späteren Zeitpunkt in dieser Arbeit noch einmal aufgegriffen wird. Diese gefundenen Stabilitätskennwerte konnten in einer weiteren Untersuchung von Howard und Bray (1988) nicht repliziert werden. In dieser Studie wurden Führungskräfte über Jahrzehnte ihres Berufslebens hinweg begleitet. Zu mehreren Zeitpunkten wurden Verhaltensstichproben erhoben, sodass Daten über einen Zeitraum von 20 Jahren verfügbar waren. Damit liegt neben einer beeindruckenden Studie von der Validität psychologischer Auswahlverfahren auch eine Untersuchung zur Stabilität und Veränderung berufsrelevanter Persönlichkeitseigenschaften vor. Die Autoren belegten Veränderungen hinsichtlich unterschiedlicher Faktoren, wie z.B. Ehrgeiz, Leistungsmotivation und das Interesse an Macht und Führung. Darüber hinaus belegten die Autoren, dass unterschiedliche

²¹ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 23 f.

²² James, 1981, S. 125 – 126 (Erstveröffentlichung des Buches 1890)

Muster von Veränderungen mit unterschiedlichen Karrierewegen in Zusammenhang stehen. Die derzeitige Befundlage legt also nahe, dass sehr breite Verhaltenstendenzen (z.B. die Faktoren des Big-Five-Modells) relativ stabil über die Zeit bleiben, während bei einigen enger definierten Dimensionen jedoch deutliche Veränderungen auftreten können. Neben den von Howard und Bray beschriebenen kontinuierlichen Veränderungen können auch abrupte und dramatische Persönlichkeitsveränderungen stattfinden, denen meist ein bestimmtes kritisches Lebensereignis vorausgeht. Diese Veränderungen werden in der vorliegenden Arbeit nicht näher betrachtet.²³

Hinsichtlich berufseignungsdiagnostischer Fragestellungen ist trotz des Potentials für Veränderungen, von langfristiger Stabilität auszugehen. Spezifischere Dimensionen verändern sich einerseits durch Reifungsvorgänge und andererseits in Abhängigkeit von bestimmten Karrierewegen. Dies ist jedoch noch zu wenig untersucht, sodass die Veränderungsmuster den modellhaft diagnostischen Entscheidungsfindungsprozessen noch nicht als Basis dienen können.²⁴

Hinsichtlich praktischer Fragestellungen ist eher die Validität besonders entscheidend. Kann auch ohne nachgewiesene Stabilität der Indikatoren eine gute Vorhersage getroffen werden, ist dies eine ausreichende Rechtfertigung für den Einsatz von Persönlichkeitsverfahren oder anderen diagnostischen Instrumenten.²⁵

²³ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 37 f.

²⁴ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 39 - 40

²⁵ vgl. ebenda, zit.n. Wottawa/ Hossiep, 1987

2.2.3 Einordnung von Persönlichkeitsverfahren

Im Rahmen jeder Eignungsuntersuchung stehen grundsätzlich drei mögliche Zugänge zu diagnostisch relevanten Informationen zur Verfügung. Diese sind in der folgenden Abbildung dargestellt:

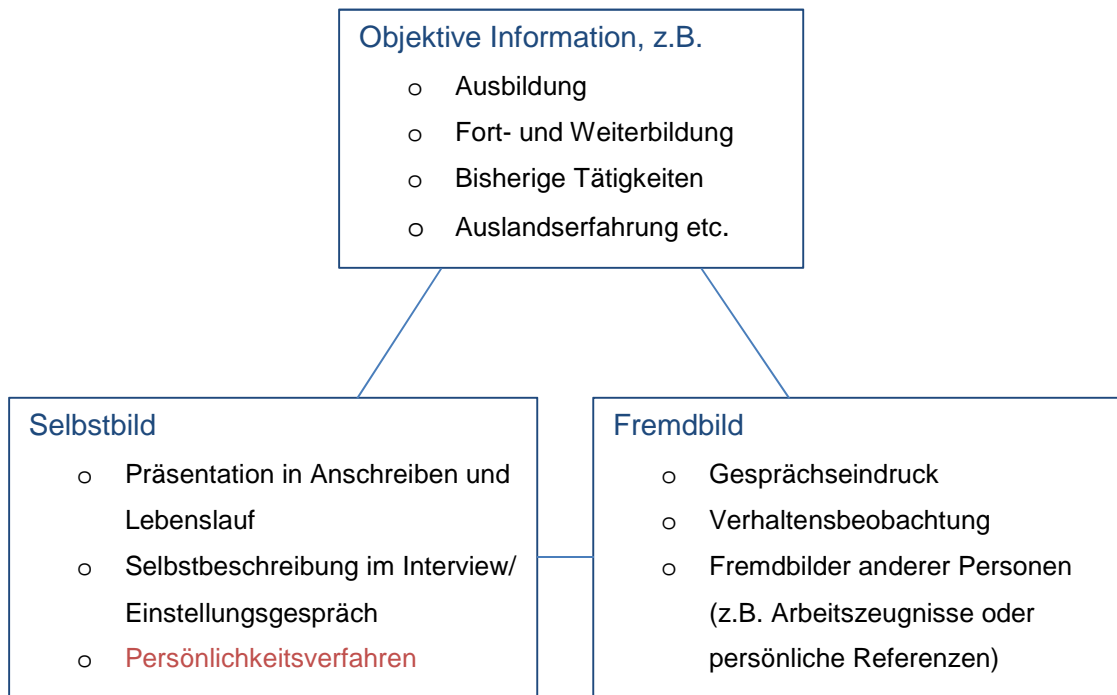


Abb. 1: Die drei Quellen diagnostischer Information (nach Hossiep/ Paschen/ Mühlhaus, 2000, S.53)

Jede dieser drei Informationsquellen kann wichtige Aspekte zur Unterstützung der Entscheidungsfindung beitragen. Im Idealfall sollte deswegen keiner dieser Zugänge vernachlässigt werden. Persönlichkeitsverfahren liefern Hinweise bezüglich des Selbstbildes des Kandidaten, wobei diese Informationen Querschnittscharakter haben, d.h. sie beziehen sich auf das augenblickliche Selbstbild. Längsschnittinformationen dagegen beziehen sich auf Beobachtungen und Einschätzungen über einen längeren Zeitraum.²⁶

²⁶ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 51 ff.

Die Klassifikation eignungsdiagnostischer Verfahren kann nach unterschiedlichen Gesichtspunkten erfolgen, z.B. nach:

- erfassten Merkmalen: kognitiv vs. nicht-kognitiv
- Art der Durchführung: schriftlich oder mündlich
- dem Standardisierungsgrad
- dem Prognosekonzept: eigenschaftsorientiert oder simulationsorientiert.²⁷

Persönlichkeitsverfahren sind hier den nicht-kognitiven Verfahren zuzuordnen, die meistens in Form von Fragebögen schriftlich durchgeführt werden.

Persönlichkeitsbeschreibende Verfahren sollen typisches Verhalten messen, wohingegen Leistungstests maximales Verhalten erfassen. Zu den Persönlichkeitsverfahren, die nicht-leistungsmäßige Eigenschaften messen sollen, gehören Persönlichkeits-Struktur-Verfahren, spezifische Persönlichkeitsverfahren (z.B. Empathie oder Kontrollüberzeugungen), Typologien, aber auch, in einer erweiterten Auffassung, Verständnis-, Interessen- und Motivationsverfahren. Verfahren zur sozialen Kompetenz nehmen dabei eine Zwischenposition ein.²⁸

2.3 Notwendigkeit von Persönlichkeitsverfahren in der psychologischen Berufseignungsdiagnostik

Obwohl Deutschland insgesamt noch immer ein Entwicklungsland bezüglich der Nutzung von Testverfahren im Personalmanagement darstellt (siehe Abschnitt 2.5), werden außer den reinen Leistungstests inzwischen vermehrt auch Persönlichkeitsverfahren eingesetzt, um nicht-kognitive Eigenschaften, die die Arbeitsleistung u.a. mitbedingen können, zu erfassen. Dies ist eine erfreuliche Entwicklung, sind doch Persönlichkeitsmerkmale für die Prognose der beruflichen Leistung kaum weniger wichtig als Leistungsmerkmale. Dies wurde bereits durch Forschungen gezeigt, die einen beachtlichen Zusammenhang zwischen den sog. „big five“, den fünf großen Persönlichkeitsfaktoren (Extraversion, Emotionale Stabilität, Offenheit für neue Erfahrungen, Gewissenhaftigkeit und Verträglichkeit) und beruflichem Erfolg nachgewiesen haben.²⁹ Auch die nachfolgende Abbildung zeigt,

²⁷ vgl. Schuler/ Funke, 1995, S. 235 ff.

²⁸ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. XVI

²⁹ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. XVII

dass kognitive Fähigkeiten nicht allein dafür verantwortlich sind, ob man im Berufsleben erfolgreich ist oder nicht.

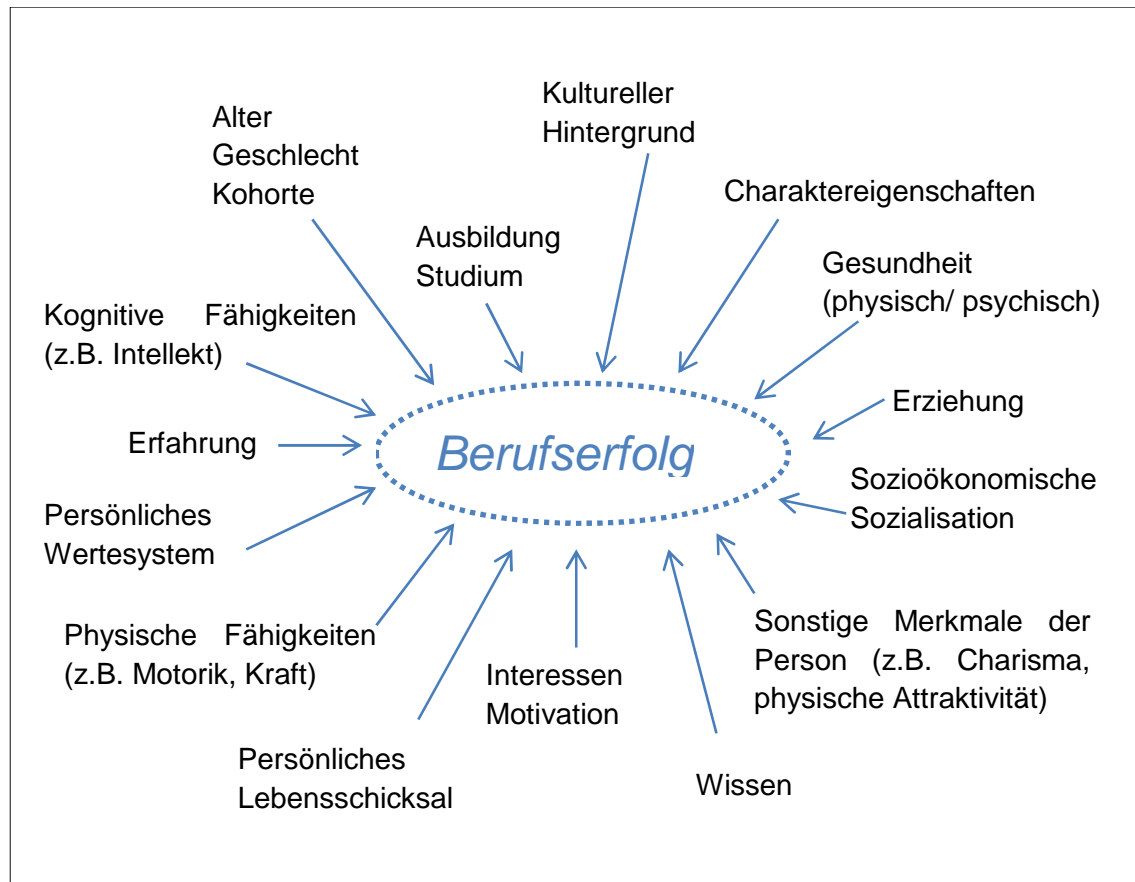


Abb. 2: Determinanten des Berufserfolges (adaptiert nach Hossiep, 2007)

Experten schätzen, dass über 90% der Beschäftigungsverhältnisse nicht aufgrund der fachlichen Kompetenz eines Mitarbeiters beendet werden, sondern eher wegen Unstimmigkeiten von Merkmalen seiner Persönlichkeit mit den Anforderungsbedingungen der Position.³⁰

³⁰ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. XVII

Dies bestätigt auch eine Studie von Kienbaum aus dem Jahr 2005, in der 230 Unternehmen an einer Befragung teilnahmen. Unter anderem wurde nach den Gründen für das Scheitern von High Potentials in Unternehmen gefragt. Die Ergebnisse sind in der folgenden Abbildung dargestellt:

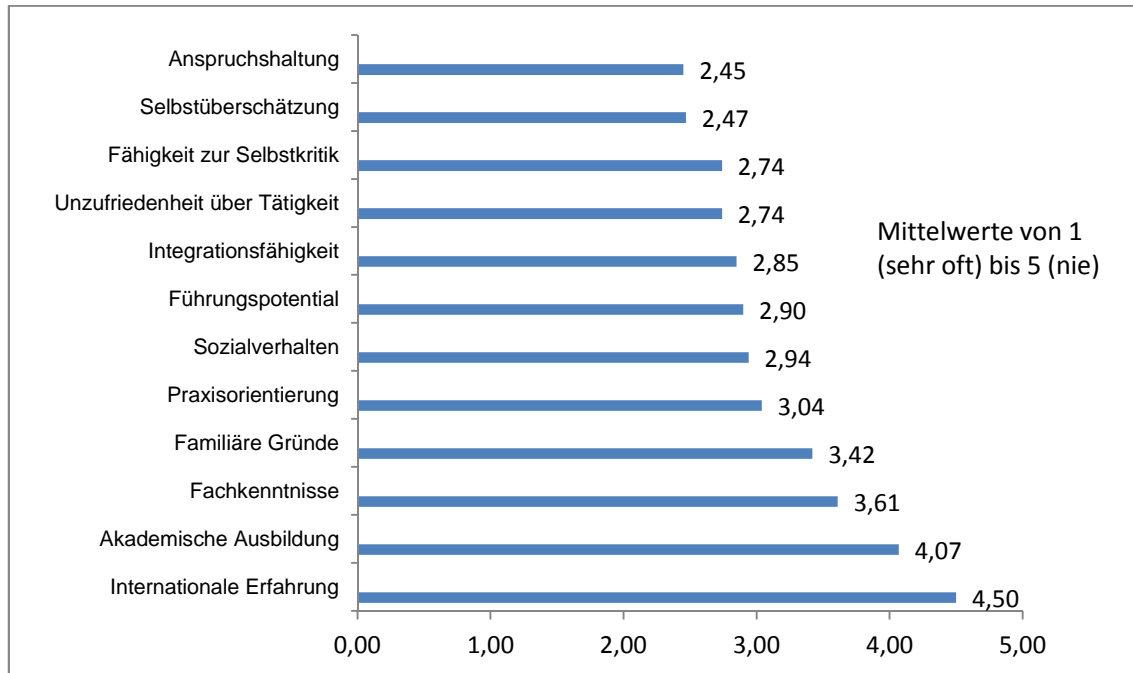


Abb. 3: Woran scheitern High Potentials im Unternehmen? (Quelle: Kienbaum, 2005)

Auch hier ist ersichtlich, dass Fachkenntnisse und die akademische Ausbildung nicht die hauptsächlichen Gründe für ein Scheitern darstellen. Meist liegen die Gründe im Bereich von persönlichen Eigenschaften, wie Selbstüberschätzung oder mangelnder Fähigkeit zur Selbstkritik.

Die tatsächlichen Effekte und Konsequenzen, die sich für ein Unternehmen hieraus ergeben, können, werden immer noch vielfach unterschätzt. Persönlichkeitsverfahren werden aber nicht nur wegen der Unterschätzung der Bedeutung von Merkmalen aus dem Persönlichkeitsbereich noch viel zu selten eingesetzt, sondern auch wegen einer eher kritischen Haltung und ideologischen Ablehnung gegenüber diesen Verfahren: Die wesentlichen Argumente der Kritiker von Persönlichkeitsverfahren sind folgende:

- In die Privatsphäre des Mitarbeiters oder Kandidaten wird unberechtigt eingedrungen
- Es besteht die Möglichkeit des Missbrauchs bzw. der Weitergabe von Daten.
- Die erhaltenen Persönlichkeitsbilder bleiben unvollständig.

- Es besteht häufig kein offensichtlicher Bezug zum Arbeitsinhalt.
- Situative Besonderheiten der späteren Arbeit werden nicht berücksichtigt.
- Häufig werden die Verfahren von Personen angewendet, die keinerlei Vorbildung in der sachgerechten Anwendung und Interpretation dieser diagnostischen Instrumente haben.³¹

Diese Argumente treffen aber nicht nur auf Persönlichkeitsverfahren zu, sondern auch für die meisten anderen Verfahren der psychologischen Berufseignungsdiagnostik. Die meisten von ihnen können entkräftet werden. Rechtliche Rahmenbedingungen z.B. sorgen dafür, dass nicht unberechtigt in die Privatsphäre des Menschen eingedrungen wird. Vertieft wird diese Problematik noch einmal in Abschnitt 2.7, in dem rechtliche und ethische Aspekte der Berufseignungsdiagnostik dargestellt werden. Ein weiteres Argument gegen den Einsatz von Persönlichkeitsverfahren stellt das Verfälschen der Ergebnisse in eine positiv angenommene Richtung dar, so dass die erhaltenen Informationen unbrauchbar sind. Man spricht von „Sozialer Erwünschtheit“. Dem ist entgegenzuhalten, dass die Verfälschungs-wahrscheinlichkeit vernachlässigbar gering ist, wenn die Kandidaten mit einer grundsätzlichen Überprüfbarkeit ihrer Angaben, z.B. durch nachfolgende Interviews oder Assessment Center o.ä. rechnen müssen.³² Dieses Problem wird zu einem späteren Zeitpunkt in der Arbeit noch einmal genauer diskutiert (siehe Abschnitt 2.8).

Festzustellen ist, dass sich hierzulande die restriktive Haltung gegenüber der psychologischen Berufseignungsdiagnostik allgemein und den Persönlichkeitsverfahren im Besonderen doch allmählich ändert. Dies ist darauf zurückzuführen, dass aufgrund von Lebensläufen und Zeugnissen der Kandidaten keine gültigen Aussagen mehr über die Fähigkeiten und Fertigkeiten getroffen werden können. Sie verlieren immer mehr an Aussagekraft. Schulnoten aus Zeugnissen geben fast ausschließlich einen Einblick in die intellektuellen Fähigkeiten des Kandidaten. Bei der Betrachtung einzelner Noten stößt man auf ein grundlegendes Problem. Die Benotung schwankt stark nach Schulform, Bundesland und einzelner Schulen. Außerdem richtet sich die Notengebung an das uneinheitliche Anspruchsniveau der Lehrkräfte, das sich in erster Linie an der Leistung des Einzelnen in der Relation zu seinen Mitschülern ergibt und sich nicht an der absoluten Leistung orientiert. Ein Schüler in einer leistungsstarken Klasse muss mitunter mehr leisten, um eine gute Note zu erhalten als in einer vergleichsweise leistungsschwachen Klasse. Die Schulnoten lassen einen eher strengen oder liberalen Bewertungsmaßstab nicht

³¹ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. XVII

³² vgl. ebenda

erkennen und machen einen direkten Vergleich der Kandidaten eigentlich unmöglich.³³ Ein ähnliches Problem tritt auch bei Hochschulabsolventen auf. Es ist schwierig die verschiedenen Abschlüsse untereinander zu vergleichen. Was ist ein Bachelor- oder Masterabschluss heute wert? Wo ordnet sich das traditionelle Diplom ein? Und ist ein Masterabschluss einer bayrischen Hochschule mehr wert als der einer sächsischen Hochschule? Diesen und ähnliche Fragen müssen sich immer mehr Unternehmen stellen. Die psychologische Berufseignungsdiagnostik setzt genau an diesem Punkt an und auch Persönlichkeitsverfahren spielen dabei eine nicht zu unterschätzende Rolle.

Die tatsächliche Effektivität des Einsatzes von Persönlichkeitsverfahren lässt sich wissenschaftlich insbesondere mit der prognostischen Validität (siehe Abschnitt 2.6.2) belegen. Einschlägige Untersuchungen weisen eindeutig darauf hin, dass die ergänzende Untersuchung bestimmter Persönlichkeitsdimensionen mit wissenschaftlicher Methodik einen zusätzlichen Nutzen erbringt. Die Gültigkeit der diagnostischen Instrumente und den damit verbundenen Entscheidungen kann durch den Einsatz von Persönlichkeitsverfahren erhöht werden. Belege hierfür finden sich vor allem in Metaanalysen, wie z.B. der von Schmidt und Hunter (1977, 1990). Sie entwickelten eine Methode, mit der die Generalisierbarkeit von Validitätskennwerten über verschiedene Situationen überprüft werden konnte. Durch statistische Korrekturen der Einzeluntersuchungen u.a. wurde es möglich, größere Anzahlen von Studien zusammenzufassen und daraus Aussagen zur Gültigkeit der untersuchten Persönlichkeitsmerkmale abzuleiten.³⁴ Die entwickelte Methodik, auch wenn sie verschiedenen Limitierungen unterliegt, war für die persönlichkeits-psychologische Forschung ein enormer Fortschritt. Es wurde erstmals überzeugend gezeigt, dass bestimmte Persönlichkeitsmerkmale über unterschiedliche Regionen, Organisationen und Jahre hinweg einen substantiellen Zusammenhang mit Kriterien des Berufserfolges aufweisen. Wesentliche Erkenntnisse aus diesen Metaanalysen können vertiefend bei Hossiep und Mülhhaus (2005) nachgelesen werden.³⁵ Die Bedeutung von Persönlichkeitsverfahren wird noch unterstrichen, wenn man Kombinationsstudien hinzuzieht (vgl. z.B. Schmidt & Hunter, 1998). Es kann klar gezeigt werden, dass Persönlichkeitsverfahren einen Beitrag leisten, der die anderen Verfahren (z.B. Interviews, Intelligenztests) ergänzt und die gesamte Validität des Prozesses verbessert.³⁶

³³ vgl. Kanning/ Pöttker/ Klinge, 2008, S.100 f.

³⁴ vgl. Hossiep, 2000, S. 80 ff.

³⁵ vgl. Hossiep/ Mülhhaus, 2005, S. 107

³⁶ vgl. ebenda

Zusammenfassend ist zu sagen, dass Persönlichkeitsverfahren in ein diagnostisches Gesamtkonzept eingearbeitet werden sollten, denn sie sind als alleinige Grundlage der Entscheidungsfindung oder als Selektionsinstrument nicht geeignet. Sie sind eher als ein nützliches Werkzeug zu verstehen, mit dem gewisse Facetten des Selbstbildes der Kandidaten standardisiert erhoben werden können. Die Ergebnisse eines Persönlichkeitsverfahrens sollten Gegenstand eines Gespräches mit dem (potentiellen) Mitarbeiter werden. Erst dann können sie einen nützlichen Beitrag zu einem sich formenden Gesamtbild des Kandidaten liefern.³⁷

Folgerichtig sollen Persönlichkeitsverfahren also in einem eignungsdiagnostischen Gesamtkonzept nie das einzige, vielmehr nur ein diagnostisches Instrument neben anderen sein.

2.4 Personalpolitische Ausgangslage

Die Personalpolitik stellt einen wichtigen Bestandteil der allgemeinen Unternehmenspolitik dar. Sie ist eng mit anderen Politikbereichen in der Unternehmung, wie z.B. der Absatz-, Produktions-, Beschaffungs-, Investitions- oder Finanzpolitik verknüpft. Nach Easton versteht man unter der Personalpolitik die „Gesamtheit der autoritativen Entscheidungsprozesse über die Allokation von materiellen und immateriellen Werten einer Unternehmung an das Personal“.³⁸ In diesem Zusammenhang kommt der Eignungsdiagnostik eine spezifische Funktion zu, die vorwiegend instrumenteller Natur ist. Sie bildet die Basis für die Verteilung, Zuordnung bzw. Besetzung von Positionen mit dem Anspruch der Berechenbarkeit und Nachvollziehbarkeit, der Zweckmäßigkeit und Sinnhaftigkeit.³⁹

Die genannten Maßstäbe gelten ebenso für die Grundsätze der Personalpolitik, u.a. für die Führung und Zusammenarbeit, die Personalauswahl und Personalentwicklung und für die Weiterbildung und Information. Dabei ist der Grundsatz der Leistung und Gleichbehandlung besonders zu berücksichtigen. Für die Anwendung eignungsdiagnostischer Verfahren in Organisationen oder Institutionen sind Leistungsfähigkeit und Leistungswillen deshalb wichtige Kriterien. Für das individuelle Leistungsverhalten hat die Zielsetzung wiederum eine wichtige Orientierungsfunktion.⁴⁰

³⁷ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 54

³⁸ vgl. Wagner, 1995, S. 43, zit. n. Engelhard, 1984

³⁹ vgl. Wagner, 1995, S. 43

⁴⁰ vgl. Kleinbeck, 1991, S. 41 f.

Die erwähnten Grundsätze der Personalpolitik bilden, in Form von personalpolitischen Strategien und Handlungsempfehlungen, die Rahmenbedingungen für Art, Umfang und Intensität eignungsdiagnostischer Methoden und Verfahren. Sie können dabei sowohl Chance als auch Restriktion darstellen.⁴¹

Die personalpolitischen Prinzipien der Personalbeschaffung, -förderung und -entwicklung sind eng verbunden mit der Art und Weise der strategischen Einordnung von Personalbesetzungen. Es stellt sich die Frage, ob und in welchem Maße strategische Überlegungen in personaler Hinsicht überhaupt vorhanden sind. Es gibt Unternehmen, in denen Personalbeschaffung eher zufällig und fallweise erfolgt, wenn unvermittelt eine personelle Lücke zu füllen ist. In anderen Unternehmen wiederum findet man ein langfristig orientiertes Management humaner Ressourcen, verbunden mit der Harmonisation ökonomischer und sozialer Interessen. Darüber hinaus ist anzunehmen, dass für unterschiedliche Entwicklungsstufen eines Unternehmens unterschiedliche Personalstrategien angezeigt sind.

Ist die Personalpolitik in einem Unternehmen stark verankert, können im Hinblick auf eignungsdiagnostische Aspekte folgende Fragen beantwortet werden:

- Welche Kriterien sind wichtig für die Personalauswahl im Rahmen der Personalbeschaffung, -planung und -entwicklung? Welche werden zukünftig wichtig sein?
- Wann, in welchen zeitlichen Abständen und bei welchen Anlässen sollen eignungsdiagnostische Verfahren eingesetzt werden?⁴²

Die Mitwirkung der beteiligten Führungskräfte des Unternehmens und ihre Identifikation mit den verfolgten personalpolitischen und -strategischen Überlegungen sind dabei wichtig. Zu bedenken ist außerdem, dass Art, Umfang und Zeitpunkt der jeweils einzusetzenden eignungsdiagnostischen Verfahren nicht nur eine Frage der strategischen Systematik ist, sondern die folgenden Aspekte eine ebenso wichtige Rolle spielen:

- Welche Stellung nimmt die Eignungsdiagnostik im Unternehmen ein? Wird sie mehr oder weniger blindgläubig als quantitativ erfassbarer Baustein für weitere personalwirtschaftliche Überlegungen angesehen oder tatsächlich als

⁴¹ vgl. Kleinbeck, 1991, S. 41 f.

⁴² vgl. ebenda

Beratungsinstrument angesichts vielfältiger Eignungsformen, komplexer Persönlichkeiten und unterschiedlichster Anforderungssituationen?

- Von wem werden die eignungsdiagnostischen Untersuchungen vorgenommen?
Die Anwendung vorgeblich einfacher und überschaubarer Auswahlmethoden durch „geschulte Laien“ (Personalreferenten, Personalleiter u.a.) kann dabei ebenso ungeeignet sein wie die umfassende Anwendung durch großzügig besetzte betriebspsychologische Abteilungen oder teurer externer Berater.⁴³

Das Gespür aller Beteiligten über die Angemessenheit diagnostischer Verfahren mit ihren Einsatzmöglichkeiten und auch -grenzen ist entscheidend. Es sollte ein ausgewogenes Verhältnis zwischen traditionellen Verfahren (wie z.B. das Einstellungsinterview) und neuzeitlichen Methoden bestehen, das bereits bei der Formulierung personalpolitischer Grundsätze seine Berücksichtigung finden sollte.⁴⁴

2.5 Der Einsatz von Persönlichkeitsverfahren im internationalen Vergleich

Was den Einsatz von Persönlichkeitsverfahren angeht, ist Deutschland eher noch ein Entwicklungsland. Es liegt im internationalen Vergleich weit hinten. Nur 7 % der befragten deutschen Personalleiter gaben bei einer Untersuchung von Schuler, Frier und Kaufmann (1993) an, Persönlichkeitsverfahren in der Personalarbeit – speziell bei der Auswahl von Führungskräften – zu nutzen.

Die nachfolgende Tabelle zeigt Umfrageergebnisse für verschiedene Verfahren im europäischen Vergleich.

⁴³ vgl. Kleinbeck, 1991, S. 41 ff.

⁴⁴ vgl. ebenda

Tab. 1: Anwendung verschiedener Auswahlverfahren für Führungskräfte im europäischen Vergleich (nach Hossiep, 1996, S. 55)

	Deutschland	Frankreich	Spanien	Benelux	Großbritannien
Strukturiertes Interview	55	76	70	91	90
Persönlichkeitsverfahren	7	42	74	71	69
Leistungstest	3	35	38	6	3
Intelligenztest	2	30	55	54	35
Biographischer Fragebogen	15	40	63	44	18
Assessment Center	14	7	10	15	33
Graphologisches Gutachten	7	89	17	19	4

Anmerkung: Angaben in Prozent der befragten Unternehmen

Auch in aktuelleren Untersuchungen wie z.B. der von Baker und Cooper (1995) spielen Persönlichkeitsverfahren in Großbritannien und Frankreich eine wichtige Rolle. In der Schweiz werden von 27,8 % der Personalberatungen Persönlichkeitsverfahren zur Unterstützung bei Personalentscheidungen herangezogen (Huber, 1989). Eine Untersuchung in Neuseeland zeigte, dass auch hier der überwiegende Teil von Personalberatungen, die mit Psychologen zusammenarbeiten, Persönlichkeitsinventare einsetzt (Dakin, Nilakant & Jensen, 1994).⁴⁵

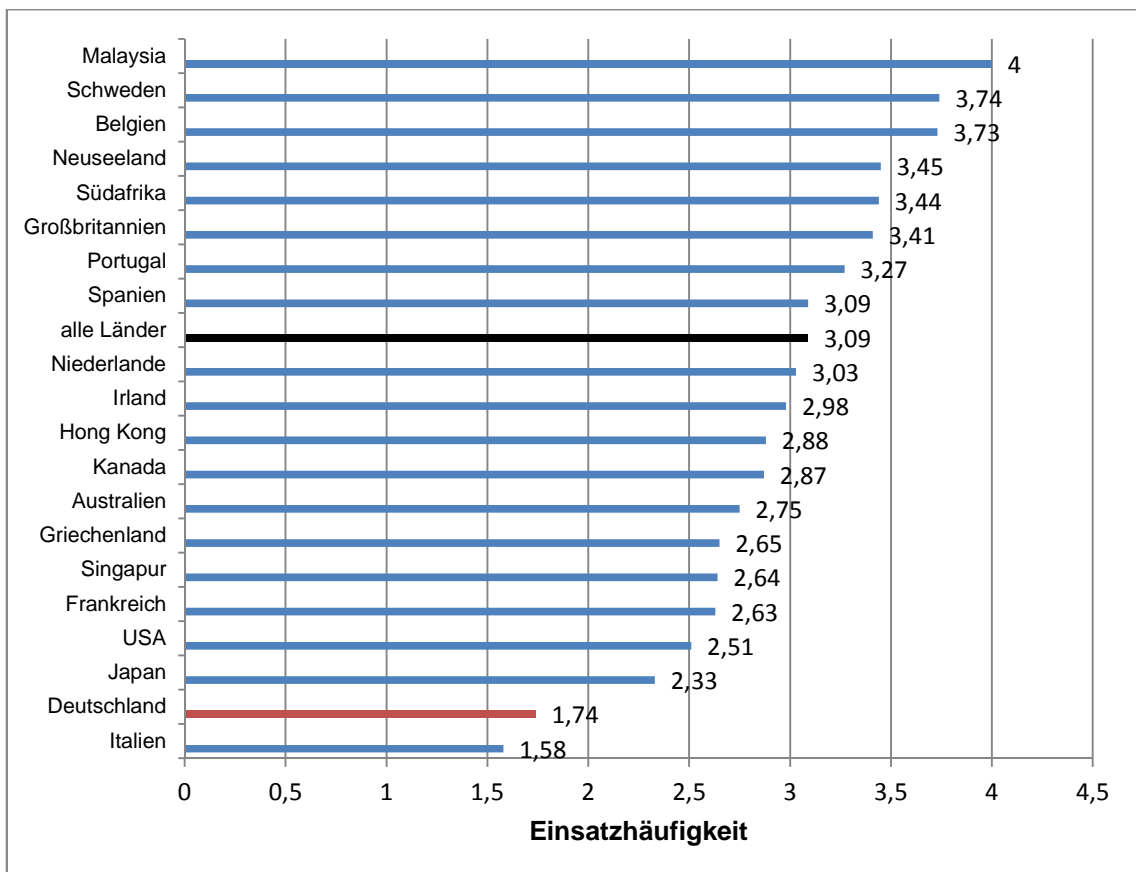
Am Häufigsten werden diese Verfahren wahrscheinlich in den Vereinigten Staaten eingesetzt. In einer umfangreichen Befragung von Arbeits- und Organisationspsychologen gaben 84,7 % der Befragten an, Persönlichkeitsverfahren bei Selektionsentscheidungen zu nutzen. Damit wurden diese in ihrer Einsatzhäufigkeit nur noch vom Interview (94,4%) übertroffen und rangieren mit den Intelligenztests sogar auf dem gleichen Rang (Ryan & Sackett, 1987).

Eine andere Umfrage von Schorr (1995) berichtet über die Verwendung von Persönlichkeitsverfahren bei deutschen Psychologen. Allerdings wurden die Ergebnisse anders aufbereitet. Somit ist ein direkter Zahlenvergleich mit den Daten aus den Vereinigten Staaten nicht möglich. Nach Schorr sind 29,1 % der Verfahren, die in Deutschland verwendet werden, Persönlichkeitsverfahren. Die übrigen 70,9 % verteilen sich dann auf andere Testverfahren. Diese Zahlen lassen leider keine Rückschlüsse auf die tatsächliche Einsatzhäufigkeit von Persönlichkeitsverfahren zu, liefern uns aber dennoch ein allgemeines Bild über die Stellung dieser in Deutschland.⁴⁶

⁴⁵ vgl. Hossiep/Paschen/ Mühlhaus, 2000, S. 44 f.

⁴⁶ vgl. ebenda

Aus den Zahlen ist abzuleiten, dass Persönlichkeitsverfahren in deutschen Unternehmen also eher seltener eingesetzt werden als in den Unternehmen anderer internationaler Länder. Dies hat sich auch in den letzten Jahren nicht einschneidend geändert. Die folgende Abbildung zeigt die Ergebnisse einer Studie von Ryan et al. Aus dem Jahr 1999:



Anmerkung: Skala von 1 (nie) bis 5 (fast immer oder immer)

Ryan et al. (1999)

Abb. 4: Einsatz von Tests oder Fragebögen bei der Personalauswahl (leitende Positionen)
(Quelle: Hossiep, 2007)

Das schwer überschaubare Testangebot und die geringe Erfahrung der meisten Unternehmen sind nur zwei mögliche Gründe für die geringe Anwendung von Persönlichkeitsverfahren. Weiterhin bestehen häufiger Vorbehalte gegen diese Verfahren, die u.a. von unprofessioneller Anwendung oder unseriösen „Psychotests“ herrühren. Was den Einsatz professioneller Auswahl- und Platzierungsinstrumente angeht, hinkt Deutschland den Ländern im anglo-amerikanischen Sprachraum immer noch deutlich hinterher. Es ist jedoch eine zunehmende Tendenz in Richtung des professionelleren Umgangs mit den unterschiedlichen Verfahren festzustellen.

2.6 Anforderungen an Testverfahren

2.6.1 Übersicht

Die Qualität eines Verfahrens wird, wie in Abbildung 5 dargestellt, durch drei zentrale Kriterien der Testgüte konstatiert: Objektivität, Reliabilität und Validität.

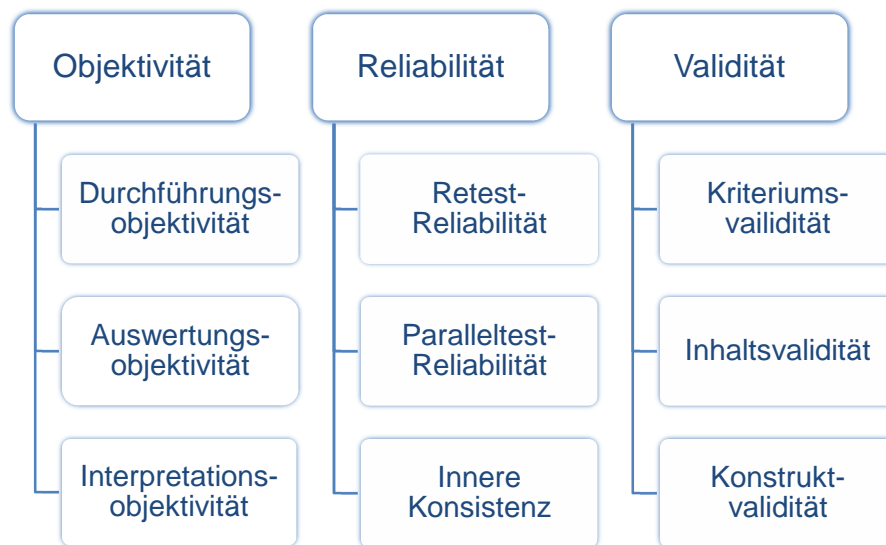


Abb. 5: Die klassischen testtheoretischen Gütekriterien (nach Skibba, 2006, S. 18)

Zusätzlich kann man Nebengütekriterien, wie Normierung, Vergleichbarkeit, Ökonomie und Nützlichkeit, zur Beurteilung eines Verfahrens heranziehen.

2.6.2 Hauptgütekriterien

Objektivität

„Ein Verfahren wird als objektiv bezeichnet, wenn die Testergebnisse sowie die aus den Ergebnissen gezogenen Schlussfolgerungen von den konkreten, jeweils variierenden Bedingungen der Testdurchführung und -auswertung unabhängig sind.“⁴⁷

Der Anspruch auf Objektivität zielt somit auf die Vergleichbarkeit von Testergebnissen verschiedener Personen. Zu den Bedingungen der Testdurchführung und -auswertung

⁴⁷ Krohne/ Hock, 2007, S. 24

gehören ganz essentiell die Personen, die den Test anwenden und auswerten. Wer den Test durchführt, auswertet oder interpretiert, soll für die Testergebnisse unwesentlich sein.⁴⁸ Dies erfordert neben der genauen Beschreibung des methodischen Vorgehens (Transparenz) auch eine gewisse Standardisierung des jeweiligen Verfahrens.⁴⁹

Üblicherweise wird bei der Objektivität zwischen der Durchführungs-, Auswertungs- und Interpretationsobjektivität unterschieden. Diese Aspekte orientieren sich an den Phasen, die im Rahmen einer diagnostischen Untersuchung zu absolvieren sind. Durch die Standardisierung des Testablaufes wird die Durchführungsobjektivität gesichert. Die Anleitungen, das Material, die Art der Testvorgabe und andere Merkmale der Testsituation, die für die Resultate der Ergebnisse verantwortlich sind, sind in standardisierten Testverfahren genau fixiert. Auf die Registrierung der in einem Test anfallenden Daten und deren Kombination oder Verrechnung zu Testergebnissen bezieht sich die Auswertungsobjektivität. Die Interpretationsobjektivität wiederum beruft sich auf weitere Schlussfolgerungen, die auf Grundlage der Testergebnisse getroffen werden. Testverfahren, die diesen Anforderungen genügen, liefern exakte Regelungen und Richtlinien für die Registrierung und Auswertung der vorliegenden Daten und die Interpretation der Testbefunde. So wird gewährleistet, dass verschiedene Personen, denen die gleichen Daten vorliegen, auch zu gleichen Ergebnissen gelangen.⁵⁰

Um die Objektivität zu sichern, müssen die Vorgaben zur Anwendung und Durchführung eines Testverfahrens entschieden befolgt werden. Eine enorme Erleichterung für die standardisierte Durchführung bieten computergestützte Verfahren. Die Durchführungsobjektivität wird durch die Automatisierung enorm erhöht. Das Problem der Auswertungsobjektivität entfällt fast gänzlich, da die Daten weder manuell registriert noch manuell zu Testwerten verrechnet werden müssen. Auch die Interpretationsobjektivität wird durch Interpretationshinweise in den Handbüchern und den Visualisierungen durch die Software stark positiv beeinflusst. Dennoch müssen auch Testleiter solcher computergestützter Verfahren professionelle Anwender oder zumindest entsprechend geschult sein.

⁴⁸ vgl. Krohne/ Hock, 2007, S.24

⁴⁹ vgl. Bortz/ Döring, 2006, S. 326

⁵⁰ vgl. ebenda

Reliabilität

Aus praktischen Gründen betrachten Testverfahren immer nur bestimmte Ausschnitte aus dem interessierenden Verhaltensbereich: Im Rahmen einer Führerscheinprüfung kann das Rückwärts-Einparken nicht an jeder erdenklichen Parklücke geprüft werden, die im Alltag eines Fahrzeugführers frei sein könnten. Das wäre tatsächlich nicht umsetzbar. Man ist also darauf angewiesen, Verhaltensstichproben zu ziehen. Diese Stichproben sind zwangsläufig limitiert und die auf ihrer Grundlage gezogenen Schlüsse nicht vollkommen fehlerfrei.⁵¹

„Die Präzision, mit der ein Test das von ihm erfasste Merkmal misst, wird als Reliabilität oder Zuverlässigkeit bezeichnet.“⁵²

Wird ein reliabler Test unter identischen Ausgangsbedingungen wiederholt, führt dieser zu reproduzierbaren Ergebnissen. Die Reliabilität ist somit neben der Objektivität ein zweites zentrales Gütekriterium von Messungen bzw. Testverfahren. Angegeben wird die Zuverlässigkeit eines Messverfahrens durch einen mathematischen Kennwert, dem sog. Reliabilitätskoeffizienten. Der Wert dieses Koeffizienten liegt zwischen 0 und 1. Demzufolge ist ein Verfahren mit einem Koeffizienten von 0 in keiner Weise zuverlässig. Dagegen ist das Ergebnis bei einem Reliabilitätswert von 1 durch keine Fehler verunreinigt und die Messung perfekt. Tatsächlich wird man einen Reliabilitätskoeffizienten von 1 nicht realisieren können, denn in der Realität treten immer Messfehler auf.⁵³

Zur Bestimmung der Reliabilität eines Verfahrens kann man zwischen den Methoden der Testwiederholung, dem Paralleltest, der Testhalbierung u.a. unterscheiden:

Die Testwiederholung (Retest-Reliabilität) sieht die zweifache Durchführung desselben Verfahrens durch dieselben Testpersonen in einem zeitlichen Abstand vor. Je besser der Zweittest die Ergebnisse des Ersttests reproduziert, desto höher ist die Messgenauigkeit oder Reliabilität des Verfahrens.

Beim Paralleltest absolvieren die Testpersonen einen Test in zwei Formen. Stimmen die Ergebnisse beider Formen in hohem Maße überein, erweist sich der Test als reliabel.

⁵¹ vgl. Krohne/ Hock, 2007, S. 25

⁵² Krohne/ Hock, 2007, S. 25

⁵³ vgl. Kanning/ Pöttker/ Klinge, 2008, S.29

Mit der Methode der Testhalbierung, dem sog. split-half-Verfahren, wird ein Test retrospektiv in zwei gleichwertige, getrennt auswertbare Aufgabengruppen unterteilt. Als Maß für die Messgenauigkeit wird die Korrelation der Ergebnisse der beiden Testhälften herangezogen.⁵⁴

Validität

Die Validität ist das wichtigste Merkmal, um die Güte eines Tests zu beschreiben.

„In Tests, Prüfungen usw. werden quantitative oder klassifizierende Aussagen über einen Erlebens- und Verhaltensbereich gemacht. Mit diesen Aussagen wird das anvisierte Merkmal mehr oder weniger gut getroffen. Das Ausmaß, in dem die Testergebnisse das zu erfassende Merkmal treffen, heißt Validität oder Gültigkeit: Ein Test ist in dem Maß valide, als er das misst, was er messen soll.“⁵⁵

Valide Tests geben uns die Möglichkeit, von den Testergebnissen auf das interessierende Merkmal zu schließen. In vielen Testverfahren oder testähnlichen Verfahren soll künftiges Verhalten vorausgesagt werden: Kann der Führerscheinprüfling zukünftig alltägliche Verkehrssituationen sicher bewältigen ohne sich und andere zu gefährden? Die Genauigkeit solcher Vorhersagen stellt einen zentralen Teilaspekt der Validität dar. Sie wird als prädiktive oder prognostische Validität bezeichnet. Um die Validität festzustellen und zu sichern, muss empirisch nachgewiesen werden, was ein Test erfasst und was nicht.⁵⁶

Da Validität ein sehr umfassendes Konzept ist, existiert im Unterschied zur Reliabilität im allgemein kein einzelner Kennwert für die Validität eines Verfahrens. Traditionell werden drei Validitätsarten unterschieden, um den Anwendungsbereich des Konzeptes zu umschreiben, nämlich Inhaltsvalidität, Kriteriumsvalidität und Konstruktvalidität. Die Unterscheidung bezieht sich dabei auf die verschiedenen Quellen, aus denen Information über die Bedeutung der Testwerte gewonnen werden kann. Die Beurteilung der Inhaltsvalidität stützt sich primär auf testinterne Informationen. Informationen hinsichtlich der Kriteriumsvalidität liefern dagegen empirische Untersuchungen über externe Korrelate der Testwerte. Für die Bewertung der

⁵⁴ vgl. Weuster, 2004, S. 15

⁵⁵ Krohne/ Hock, 2007, S. 25

⁵⁶ vgl. Krohne/ Hock, 2007, S. 26

Konstruktvalidität werden sowohl testinterne als auch testexterne Informationen genutzt.⁵⁷

➤ Inhaltsvalidität:

Die Bedeutung von Testergebnissen ist offensichtlich von der Formulierung und Zusammenstellung der Items abhängig. Ein Verfahren ist dann inhaltsvalide, wenn seine Items für den zu messenden Erlebens- und Verhaltensbereich relevant und in ihrer Gesamtheit repräsentativ sind.⁵⁸ Die Höhe der Inhaltsvalidität eines Tests kann nicht numerisch bestimmt werden. Sie beruht allein auf rein subjektiven Einschätzungen.⁵⁹ Aufgrund dessen wurde die Inhaltsvalidität in älteren Darstellungen der Diagnostik als weniger wichtig, manchmal sogar als irrelevant eingestuft. Dennoch sollte man sich nicht nur auf empirische Zusammenhänge verlassen und inhaltliche Überlegungen nicht vernachlässigen.⁶⁰

➤ Kriteriumsvalidität:

Um die Kriteriumsvalidität zu bestimmen, werden Zusammenhänge zwischen Test- und Kriteriumsvariablen überprüft. Ein Verfahren ist valide hinsichtlich eines Kriteriums, wenn der Zusammenhang stark genug ist, um auf Basis der Testwerte praktisch brauchbare Aussagen hinsichtlich der Ausprägung des Kriteriums zu machen.⁶¹

Definiert ist die Kriteriumsvalidität demnach als Korrelation zwischen den Testwerten und einem Beobachtungssachverhalt. Dieser kann entweder nahezu gleichzeitig mit dem betreffenden Test oder erst zu einem späteren Zeitpunkt gemessen werden. Man spricht von konkurrenter oder prädiktiver Kriteriumsvalidität. Die konkurrente Validität (auch Übereinstimmungsvalidität) eines Studieneignungstests könnte man z.B. erfassen, indem man die Testwerte mit Indikatoren des derzeitigen Studienerfolges, wie den aktuellen Klausurnoten u.a., korreliert. Bringt man die Testwerte vor Studieneintritt mit Maßen des späteren Studienerfolges in Zusammenhang, bestimmt man die prädiktive Validität (Vorhersagevalidität) des Tests. Die Bewährung eines

⁵⁷ vgl. Krohne/ Hock, 2007, S. 26

⁵⁸ vgl. ebenda

⁵⁹ vgl. Bortz/ Döring, 2006, S. 200

⁶⁰ vgl. Krohne/ Hock, S. 66 f.

⁶¹ vgl. ebenda, S. 68

Verfahrens an externen Kriterien, gilt als zentraler Untersuchungsaspekt für dessen Validität.⁶²

Die Kriteriumsvalidierung ist in ihrem Anwendungsbereich dahingehend stark eingeschränkt, dass vielfach kein adäquates Außenkriterium benannt werden kann. Welches objektiv beobachtbare Außenkriterium ist z.B. indikativ für das Konstrukt Neurotizismus? Die Höhe der Korrelation zwischen einem Test und einem Kriterium ist weiterhin auch eine Frage der Reliabilität der Kriteriumsmessung. Mängel in der Reliabilität drücken die Korrelationen im Allgemeinen nach unten.⁶³

➤ Konstruktvalidität

Da die Kriteriumsvalidierung nur bei geeigneten Außenkriterien sinnvoll ist und die Inhaltsvalidität keinen objektivierbaren Kennwert darstellt, kommt der Konstruktvalidität oft besondere Bedeutung zu.⁶⁴

Ein Test ist dann konstruktvalid, wenn aus dem zu messenden Zielkonstrukt Hypothesen ableitbar sind, die mithilfe der Testwerte bestätigt werden können. Man formuliert in diesem Fall, anstatt eines einzelnen Außenkriteriums, ein ganzes Netz von Hypothesen über das Konstrukt und seine Relationen zu anderen manifesten und latenten Variablen. Fallen die Testwerte so aus, wie es zuvor aus der Theorie und Empirie mittels Hypothesen abgeleitet wurde, kann dies als Hinweis für die Konstruktvalidität angesehen werden. Eine solche Konstruktvalidierung ist umso überzeugender, je mehr gut gesicherte Hypothesen ihre Überprüfung bestehen.⁶⁵

Der Begriff der Konstruktvalidität ist so weit gefasst, dass er neben der Kriteriums- und Inhaltsvalidität, auch Teile der Reliabilität als spezielle Aspekte beinhaltet. Des Weiteren wurden Konzepte eingeführt, die noch über diese Aspekte hinausgehen. Diese Konzepte kann man u.a. bei Bortz und Döring (2006) nachlesen.

⁶² vgl. Krohne/ Hock, S. 68

⁶³ vgl. Bortz/ Döring, 2006, S. 200 f.

⁶⁴ vgl. ebenda

⁶⁵ vgl. ebenda

2.6.3 Der Zusammenhang von Objektivität, Reliabilität und Validität

Zwischen den drei klassischen Gütekriterien besteht ein enger Zusammenhang. Zum Teil bedingen sie sich gegenseitig. Die Objektivität stellt generell eine notwendige, aber keine hinreichende Bedingung für die Erfüllung der Reliabilität dar. Letztere wiederum ist eine notwendige aber keine hinreichende Bedingung für Validität. Hohe Reliabilität kann also durchaus mit niedriger Validität einhergehen. Oft passiert das, wenn ein Test nur einen Teilaspekt des relevanten Merkmals mit Fragen oder Aufgaben abdeckt, d.h. wenn das faktisch gemessene Merkmal zu eng ist.⁶⁶

In der folgenden Abbildung werden die Interdependenzen noch einmal dargestellt, wobei zusätzlich nach den verschiedenen Aspekten der Objektivität unterschieden wird.

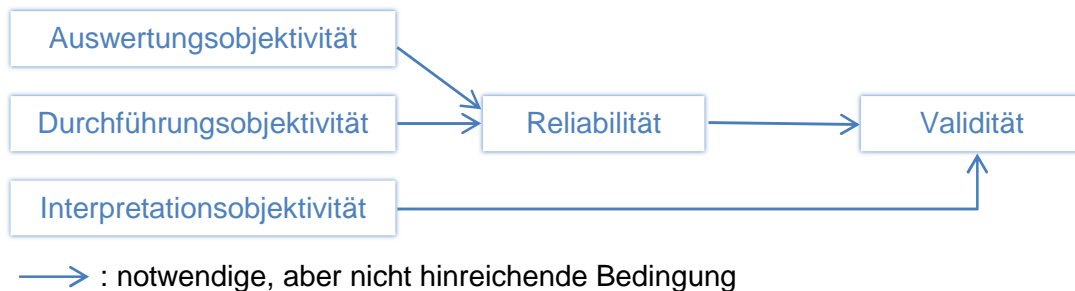


Abb. 6: Interdependenz der Gütekriterien (nach Schelten, 1997, S. 125)

Auswertungs- und Durchführungsobjektivität sind notwendige, aber keine hinreichenden Voraussetzungen für die Reliabilität eines Verfahrens. Damit ist es möglich, dass anhand eines Verfahrens objektive Werte erzielt werden können, ohne dass diese reliabel ermittelt worden sind. Diese beiden Objektivitätsaspekte stellen indirekt auch eine Voraussetzung für die Erfüllung der Validität dar, da die Reliabilität eine notwendige, aber nicht hinreichende Bedingung für die Validität ist. Die Interpretationsobjektivität stellt unabhängig von der Reliabilität ebenfalls eine notwendige, aber nicht hinreichende für die Validität dar. Sind die anhand der Testergebnisse getroffenen Aussagen also abhängig von der Person, die diese interpretiert, können keine validen Aussagen getroffen werden. Auf der anderen Seite müssen Testergebnisse, die eine hohe Interpretationsobjektivität besitzen, nicht unbedingt zu gültigen Aussagen über die zu messende Fähigkeit führen.⁶⁷

⁶⁶ vgl. Krohne/ Hock, 2007, S. 26

⁶⁷ vgl. Schelten, 1997, S. 124 f.

2.6.4 Nebengütekriterien

Für alle Nebengütekriterien gilt, dass für sie im Gegensatz zu Validität und Reliabilität keine zahlenmäßigen Kennwerte existieren.⁶⁸

Normierung

Die Normierung greift die Frage nach dem Bezugssystem eines Verfahrens auf. Es sollten Angaben vorliegen, die als Bezugssystem für die Einordnung des individuellen Testergebnisses eines Probanden dienen können. Die Resultate eines Probanden lassen sich als unterdurchschnittlich, durchschnittlich oder überdurchschnittlich im Vergleich zu anderen bezeichnen. Der Normwert stellt somit einen Vergleichswert dar.⁶⁹ Die Normierung kann für die Gesamtpopulation eines Areals (Gesamtnormen) oder für die Population einer bestimmten sozialen Gruppe (Gruppennormen) erfolgen. Auch eine Normierung für mehrere Populationen ist möglich.⁷⁰

Die betrachteten Normen sollen in Bezug auf relevante Kriterien den Merkmalen des Probanden entsprechen. Der Proband sollte u.a. der gleichen Alters- und Bildungsgruppe zugehörig sein.⁷¹

Ein Verfahren, das zwar die Hauptgütekriterien erfüllt, aber keine Normierung besitzt, ist diagnostisch nur bedingt nutzbar, da die Testergebnisse allein keinen Vergleich erlauben.

Vergleichbarkeit

Vergleichbar ist ein Testverfahren dann, wenn mindestens eine Paralleltestform vorhanden ist oder andere Instrumente verfügbar sind, die sich in ihrer Validität ähneln, also entsprechend dasselbe oder ein nahe verwandtes Merkmal messen.

Die Parallelform eines Tests ermöglicht einen Vergleich des Tests mit sich selbst, indem man einen bestimmten Probanden mit beiden Testformen untersucht und die Resultate vergleicht. Sind Korrelationen zwischen zwei validitätsähnlichen Tests bekannt, wird der gleiche Proband mit beiden Tests untersucht und die Ergebnisse verglichen.⁷²

⁶⁸ vgl. Lienert/ Raatz; 1998, S. 12

⁶⁹ vgl. Rentzsch/ Schütz, 2009, S. 230

⁷⁰ vgl. Lienert/ Raatz; 1998, S. 12

⁷¹ vgl. Rentzsch/ Schütz, 2009, S. 230

⁷² vgl. Lienert/ Raatz, 1998, S. 12

Ökonomie

Unter dem Gütekriterium der Ökonomie werden u.a. ein möglichst geringer Zeitaufwand bei der Durchführung und Auswertung, eine einfache Handhabung und ein geringer Materialaufwand verstanden. Auch die Durchführbarkeit als Gruppentest spielt hier eine Rolle.⁷³

Weiterhin führt die Etablierung, Durchführung und Weiterentwicklung eines eignungsdiagnostischen Gesamtkonzeptes zu Kosten. Diese sollten im Sinne der Ökonomie bei der Bewertung eines Verfahrens ebenfalls berücksichtigt werden.

Die Zielsetzung eines Modells von Brodgen (1949) sowie von Cronbach und Gleser (1965) besteht darin, den Nutzen einer diagnostischen Prozedur für eine Selektionsaufgabe in Geldeinheiten auszudrücken. Dieses BCG-Modell (nach den Namen der Autoren benannt) soll also bestimmen, wie hoch der von dem Einsatz eines Verfahrens zu erwartende Gewinn oder Verlust ausfällt. Somit lässt sich prüfen, ob ein Konzept überhaupt profitabel ist. Weiterhin können verschiedene in Frage kommende Verfahren hinsichtlich ihres relativen Nutzens miteinander verglichen werden.

Die mit dem Modell zu schätzende Größe ist der inkrementelle Nutzen eines diagnostischen Verfahrens, d.h. der Nutzenzuwachs, der sich durch die (zusätzliche) Verwendung eines Tests oder einer Testbatterie ergibt.⁷⁴

Der Nettonutzen wird üblicherweise in Geldbeträgen ausgedrückt und kann mit der nachfolgend dargestellten Formel berechnet werden.

⁷³ vgl. Rentzsch/ Schütz, 2009, S.234

⁷⁴ vgl. Krohne/ Hock, 2007, S. 195 ff.

BSG Modell

$$\Delta U = N_a \times T \times \bar{Z}_x \times r_{xy} \times SD_y - C \times N_b$$

Die Symbole bedeuten:

ΔU	Nettonutzen, Nutzenzuwachs gegenüber einer Zufallselektion
N_a	Zahl akzeptierter Bewerber
T	mittlere Verweildauer (engl. <i>tenure</i>) der Akzeptierten in der Institution
\bar{Z}_x	Mittlerer standardisierter Testwert der Akzeptierten
r_{xy}	Validität des Verfahrens in der Population der Bewerber
SD_y	Leistungsstreuung, Wert der Differenz einer Standardabweichung im Kriterium
C	Kosten des Verfahrens pro Bewerber
N_b	Zahl der Bewerber

Abb. 7: Das BSG-Modell (Quelle: Krohne/ Hock, 2007, S. 196)

Dieses Modell ermöglicht also die Berechnung des finanziellen Gewinns bzw. Verlusts, der bei der Durchführung eines bestimmten diagnostischen Konzeptes zu erwarten ist. Auch der Nutzenvergleich verschiedener Verfahren ist eine wichtige Anwendung dieses Modells. Es lässt sich abschätzen, inwieweit es lohnenswert ist ein bestehendes Konzept durch die Hinzunahme neuer Tests zu modifizieren.⁷⁵

Man geht davon aus, dass der inkrementelle Nutzen bei der zusätzlichen Anwendung von Testverfahren zunächst relativ stark steigt, dann jedoch gegen Null tendiert. Mit dem BSG-Modell ist es möglich die ökonomisch-optimale Anzahl von Verfahren festzustellen.

Cronbach und Gleser (1965) leisteten mit der Netto-Nutzen-Funktion einen eminent wichtigen Beitrag für die psychologische Diagnostik, die erst dadurch um die ökonomische Dimension erweitert wurde. In einer Zeit des zunehmenden Wettbewerbes in allen Bereichen des menschlichen Handelns wird sie wahrscheinlich auch in Zukunft an Bedeutung nicht verlieren.⁷⁶

⁷⁵ vgl. Krohne/ Hock, 2007, S.198

⁷⁶ vgl. Amelang/ Schmidt-Atzert, 2006, S. 412 f.

Nützlichkeit

Ein Verfahren ist dann nützlich, wenn es für die Praxis relevant und fair ist. Fairness bedeutet in diesem Zusammenhang, dass das gleiche Merkmal in gleicher Präzision erfasst werden sollte, ohne systematischen Verzerrungen zu unterliegen. Es muss für alle Gruppen von Testpersonen gleich valide und reliabel sein. Probanden sollten weder aufgrund ihres Geschlechts oder ihrer Herkunft benachteiligt werden.⁷⁷

Berufsbezogene Eignungsdiagnostik darf sich aber nicht darauf beschränken, lediglich die technische Qualität der verwendeten Verfahren zu beachten, denn sie ist in einen sozialen Kontext eingebettet, den es in vielerlei Weise zu berücksichtigen gilt.⁷⁸ Auf diese Rahmenbedingungen soll im Folgenden eingegangen werden.

2.7 Anforderungen an Unternehmen

2.7.1 Juristische Aspekte der Berufseignungsdiagnostik

Mit Beginn des Einsatzes psychologisch fundierter Diagnoseverfahren im Rahmen der Personalarbeit stellte sich zunächst vorrangig die Frage, ob diese Verfahren hinsichtlich ihrer Gütekriterien (siehe Kapitel 2.4 Anforderungen an Testverfahren) den traditionellen Verfahren überlegen sind oder zumindest entsprechen. Mit wachsender Anwendungsbreite und der Verfeinerung der diagnostischen Instrumente rückte seit Anfang der 60er Jahre eine andere Fragestellung in den Mittelpunkt. Die Beteiligten wurden zunehmend sensibler bezüglich des Themas der rechtlichen Zulässigkeit sowie des rechtlichen Rahmens der Personaldiagnostik. Persönlichkeitsverfahren stehen hier ganz besonders im Fokus der Diskussionen. K.H. Schmidt (1971) bemüht sich erstmals Zulässigkeitskriterien auf Basis der herrschenden deutschen Rechtslage für die Durchführung psychologischer Testverfahren im Personalbereich zu entwickeln. Es folgten Klein und Kuznik (1982). Alle drei Autoren beschäftigten sich mit der Anwendung von Tests bzw. von Testbatterien bei Bewerbern (d.h. im Anbahnungsverhältnis) oder bei Mitarbeitern der betreffenden Organisationen.⁷⁹

⁷⁷ vgl. Rentzsch/ Schütz, 2009, S. 234

⁷⁸ vgl. Schuler, 2000, S. 6

⁷⁹ vgl. Comelli, 1995, S. 108 f.

Bezüglich der gesetzlichen Vorschriften im Bereich der Berufseignungsdiagnostik, werden in Deutschland weitgehend einheitliche Ansichten vertreten. Eine umfangreiche Rechtsprechung am konkreten Fall, welche eine bessere Orientierung erlauben würde, fehlt jedoch. Es ist immer noch selten, dass Bewerber oder betroffene Mitarbeiter gegen den Arbeitgeber wegen der Verletzung gesetzlicher Vorschriften im Rahmen einer Eignungsuntersuchung klagen.⁸⁰ Ganz bestimmte Fragen der Durchführung und Interpretation von diagnostischen Verfahren im Rahmen von Eignungsuntersuchungen sind jedoch gesetzlich geregelt. Über die rechtlichen Fragen zum Einsatz von Persönlichkeitsverfahren in den Vereinigten Staaten informiert O'Meara (1994). Nachfolgend soll ein Überblick über die rechtliche Situation gegeben werden, ohne auf entsprechende Paragraphen im Einzelnen einzugehen.

K.H. Schmidt versteht unter Eignungsdiagnostik „die Sammlung von Erkenntnissen über die Eignung einer Person für genau umschriebene Aufgaben“ und demnach „darf es auch in der Eignungsuntersuchung allein darum gehen, die Begabungsstruktur und das Persönlichkeitsbild eines Bewerbers in Beziehung zu den fest umrissenen Aufgaben des neuen Arbeitsplatzes zu stellen und aus diesen speziellen Beziehungen Prognosen für die zu erwartenden Berufsleistungen abzuleiten“.⁸¹ Daraus ergeben sich folgende juristisch zu betrachtende Aspekte:

- Der diagnostische Eingriff muss gerechtfertigt sein. Hier steht der Schutz der Menschenwürde und der Persönlichkeitsrechte des Bewerbers/ Mitarbeiters dem berechtigten und schutzwerten Interesse des Arbeitgebers an Informationen über die persönliche Eignung eines Kandidaten gegenüber.
- Die Bedingungen für die Zulässigkeit der Durchführung von diagnostischen Aktivitäten, deren Auswertung und für die Erstellung und Weiterleitung der Eignungsaussage müssen gewährleistet sein.
- Die ordnungsgemäße Behandlung bzw. Aufbewahrung der im Verlauf des eignungsdiagnostischen Verfahrens angefallenen Unterlagen muss sichergestellt werden.⁸²

Grundsätzlich wird das Informationsstreben eines Arbeitgebers im Rahmen der Eignungsbeurteilung als berechtigt und schutzwert angesehen. Besteht schon ein

⁸⁰ vgl. Comelli, 1995, S. 109 f.

⁸¹ vgl. Comelli, 1995, S. 110 nach Schmidt, K.H., 1980, S. 2445

⁸² vgl. Comelli, 1995, S.110

Vertragsverhältnis, obliegen dem Arbeitgeber diverse Nebenpflichten. So muss er bei allen seinen Maßnahmen die Persönlichkeit des Mitarbeiters respektieren. Dies gilt auch schon im Anbahnungsverhältnis, d.h. auch für einen potentiellen Mitarbeiter. Von juristischer Seite als auch in den „Grundsätzen für die Anwendung psychologischer Eignungsuntersuchungen in Wirtschaft und Verwaltung“⁸³ wird gefordert, dass jeglicher Eignungsdiagnostik eine exakte und unter Anwendung wissenschaftlicher Verfahren erstellte Analyse der Arbeitsanforderungen voranzugehen hat. So wird der Bezugsrahmen, innerhalb dessen der eignungsdiagnostische Prozess abzulaufen hat, geschaffen.⁸⁴ Die Auswahl der Merkmale zur Beurteilung und Bewertung von Probanden sollte sich dabei auf jene beschränken, die für die vorgesehene Arbeit bzw. Tätigkeit erheblich sind.

Speziell den Persönlichkeitsverfahren gilt hier besonderes Augenmerk. Es dürfen nur die Persönlichkeitsmerkmale oder Anteile erfasst werden, die sich bei der Ausführung der betreffenden Arbeit unmittelbar auswirken können. Mit steigendem hierarchischem Rang einer Position nimmt die Komplexität der Anforderungen in aller Regel ebenfalls zu. Damit erweitern sich auch die Zugriffsmöglichkeiten für die Erfassung von Persönlichkeitsmerkmalen. Bezüglich der rechtlichen Zulässigkeit kann man nur schwer eine Einschätzung treffen. Dazu muss die spezielle Ausgestaltung der betreffenden Position berücksichtigt werden. Einer Erfassung sozialen Verhaltens, der Kommunikations- und Kontaktfähigkeit, der Führungsfähigkeit, der Initiative oder Risikobereitschaft steht somit nichts entgegen, wenn eine entsprechende Position zu besetzen ist.⁸⁵

Der Arbeitgeber verstößt gegen seine Fürsorgepflicht, sobald er psychologische Verfahren einsetzt, die gegen das allgemeine Persönlichkeitsrecht verstoßen. Dieses Recht ist kein geschriebenes Recht, sondern sogenanntes Richterrecht. Der Bundesgerichtshof hat das allgemeine Persönlichkeitsrecht unter Hinweis auf Artikel 1 – Menschenwürde – und Artikel 2 – Persönlichkeitsentfaltung – des Grundgesetzes entwickelt.⁸⁶ Danach liegt ein tatbestandsmäßiger Eingriff in das Persönlichkeitsrecht vor, wenn in die Individual-, Privat- oder Intimsphäre eingegriffen wird.

⁸³ erlassen von der Sektion Arbeits- und Betriebspsychologie des Berufsverbandes Deutscher Psychologen e.V., 1988)

⁸⁴ vgl. Comelli, 1995, S.110 nach Schmidt, K.H., 1980, S.2445

⁸⁵ vgl. Comelli, 1995, S.110 f.

⁸⁶ vgl. Brox, 1995, Rn. 673

Wie in der folgenden Abbildung dargestellt werden in der Rechtswissenschaft diese drei Sphären im Rahmen eines Schichtenmodells der Persönlichkeit unterschieden.⁸⁷

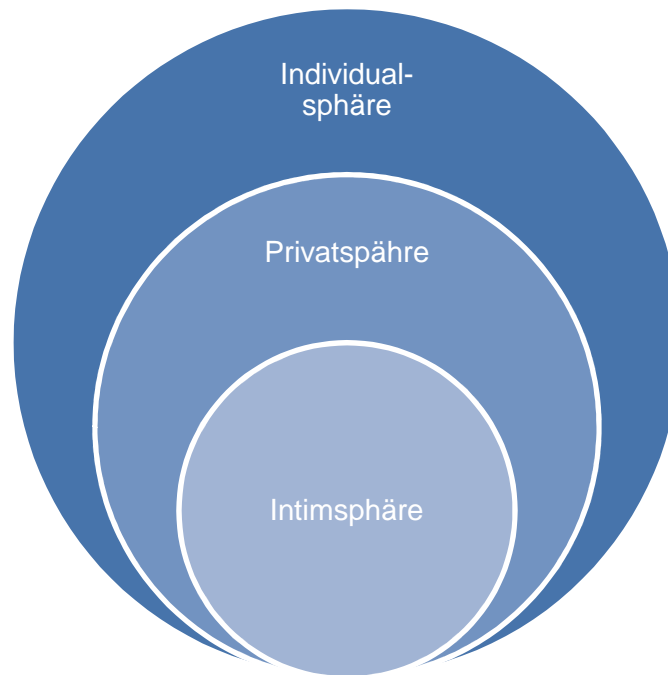


Abb. 8: Schichtenmodell der Persönlichkeit im Rahmen des allgemeinen Persönlichkeitsrechts (nach Hossiep/ Paschen/ Mühlhaus, 2000, S. 47)

Die tieferen Schichten werden vom Gesetzgeber stärker geschützt. Einige Persönlichkeitsmerkmale sind aber der Privat- oder Intimsphäre zuzuordnen. Je tiefer die Fragen der Persönlichkeitsverfahren in diese beiden Sphären eindringen, umso massiver ist die Verletzung des allgemeinen Persönlichkeitsrechts. Diese Argumentation wird vor allem von Kritikern des Einsatzes von Persönlichkeitsverfahren stark verfolgt.⁸⁸

Jede psychologische Begutachtung ist grundsätzlich mit mehreren tatbestandsmäßigen Eingriffen in das allgemeine Persönlichkeitsrecht verbunden.⁸⁹ Diese Eingriffe sind nur dann gerechtfertigt, wenn der Proband in die psychologische Eignungsuntersuchung einwilligt. Normalerweise geschieht dies durch konkludentes, d.h. schlüssiges Verhalten, indem der sich der Proband widerspruchlos dem eignungsdiagnostischen Prozess unterwirft. Diese Form der Einwilligung schließt die Durchführung und Auswertung der Untersuchung sowie die Weiterleitung des Ergebnisses mit ein, nicht aber die Aufbewahrung der Unterlagen. Erklärt der Proband

⁸⁷ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 46 f.

⁸⁸ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 47

⁸⁹ vgl. Klein, 1982, S.148

seine Einwilligung ausdrücklich, dürfen die Unterlagen verwahrt werden.⁹⁰ Eine solche Einwilligung, unabhängig davon ob sie ausdrücklich oder konkludent erteilt wurde, ist nur rechtswirksam, wenn der Proband über die Art, den Umfang und das Ziel des durchgeführten Verfahrens aufgeklärt wird. Außerdem muss eine eindeutige Beziehung zum Arbeitsplatz bestehen und das Verfahren darf nicht gegen die guten Sitten verstoßen. Der Proband muss darüber informiert werden, „was während der Untersuchung auf ihn zukommt, inwieweit das Untersuchungsergebnis an den Arbeitgeber weitergegeben wird und was mit den Untersuchungsunterlagen geschieht“.⁹¹ Bei der Weitergabe der Unterlagen ist davon auszugehen, dass der Proband nur soweit einwilligt, als diese einen Bezug zum potentiellen Arbeitsplatz haben und für die anstehende Auswahlentscheidung notwendig sind. Die Ergebnisse einer eignungsdiagnostischen Untersuchung dürfen nur an Stellen bzw. Personen des Betriebes weitergegeben werden, die für die Einstellungsentscheidung maßgebend sind. Der Testkandidat selbst kann sowohl vom Arbeitgeber als auch von dem die Untersuchung Durchführenden verlangen, über die das Ergebnis der Untersuchung unterrichtet zu werden.⁹²

Strittig ist noch immer, ob der Proband einen Vernichtungsanspruch bezüglich der Testunterlagen hat. In einigen Unternehmen und Institutionen ist es schon geübte Praxis, die Untersuchungsunterlagen nach Abschluss des diagnostischen Verfahrens zu vernichten. Dies ist dahingehend bedenklich, dass damit auch die exakte Dokumentation über Inhalt und Ablauf des durchgeführten Verfahrens vernichtet wird. Für den Fall, dass ein Bewerber eingestellt wird, darf das Gutachten bzw. das Ergebnis der eignungsdiagnostischen Untersuchungen aufbewahrt werden, da es zur Dokumentation seiner Laufbahn im Unternehmen gehört. Wird der Bewerber abgelehnt, ist sein Anspruch auf Vernichtung seiner Unterlagen sehr gut vorstellbar, da der Arbeitgeber kein berechtigtes Interesse an der Aufbewahrung haben kann.⁹³ Angesichts der vermehrten Anwendung von computergestützter psychologischer Diagnostik, kommt dem Datenschutz eine immer wichtigere Bedeutung zu.

Da das Persönlichkeitsrecht ein sog. Richterrecht ist und nicht explizit schriftlich niedergelegt wurde, werden auch in Zukunft immer wieder nicht eindeutig geklärte Fragestellungen auftreten.

⁹⁰ vgl. Comelli, 1995, S.114

⁹¹ Klein, 1982, S. 148

⁹² vgl. Comelli, 1995, S.115 f.

⁹³ vgl. ebenda

2.7.2 Ethische Aspekte der Berufseignungsdiagnostik

Die Diskussion über die ethischen Aspekte der Psychodiagnostik erlebte 1975 einen Höhepunkt mit zwei Tagungen in der Schweiz. Man setzte sich mit den Möglichkeiten und Grenzen psychologischer Erfassung auseinander. Diese Diskussion wurde in den nachfolgenden Jahren kontrovers und teilweise sehr heftig fortgeführt. Dabei ging es vor allem um die Fragen, ob Psychodiagnostik überhaupt betrieben werden sollte und – falls ja – wie sie dann zu betreiben ist. Eine sehr kritische Position beziehen Grubitzsch und Rexilius (1978). Rexilius lehnt Tests z.B. grundsätzlich ab. Er sieht in ihnen fragwürdige Messinstrumente. Die Eignungsdiagnostik wird als im Dienst des Kapitals stehend erlebt und der Vorwurf der Chancenungleichheit wird erhoben. Vertreter der anderen Seite sind z.B. Wottawa und Amelang (1985). Sie setzen sich mit dem Thema der Chancengleichheit zugunsten einer verantwortlich betriebenen Diagnostik auseinander. Aspekte der Chancengleichheit werden auch von Schuler und Funke (1987) berührt. Sie setzen sich für Fairness bei der Eignungsdiagnostik ein und verstehen darunter, dass bei der psychologischen Diagnostik Angehörige unterschiedlicher sozialer Gruppen (Frauen, Ausländer, u.a.) nicht benachteiligt werden dürfen.⁹⁴

Die betriebliche Personalarbeit wird von F.W. Schmidt (1978, 1988) der „erfassenden“ psychologischen Diagnostik zugeordnet. Dabei sollen möglichst objektive und präzise bestimmte Verhaltensmerkmale und Fähigkeiten von Personen erfasst werden. Der diagnostische Akt findet im Hinblick auf ganz bestimmte Anforderungskonstellationen statt, wobei in aller Hinsicht auch ein Vergleich mit anderen Personen beabsichtigt ist. Diese Vorgehensweise sieht Schmidt stark dem Nützlichkeitsdenken verpflichtet, denn es geht nicht um das Verständnis der ganzheitlichen Persönlichkeit.⁹⁵

Bei der Durchführung von berufseignungsdiagnostischen Verfahren steht die höchste Achtung der Person des Probanden, die strikte Einhaltung der zugesagten Vertraulichkeit und die Vorbereitung, Planung und Durchführung der Eignungsdiagnose nach dem Stand der Wissenschaft an erster Stelle.

Diagnostische Entscheidungen werden meistens unter Unsicherheit gefällt. Dies kann jedoch kein Argument gegen die Durchführung psychologischer Diagnostik sein, denn auch jede andere Entscheidungsstrategie ist fehlerbehaftet. Eine professionell betriebene Diagnostik kann den Handlungsspielraum von Individuen sogar erweitern

⁹⁴ vgl. Comelli, 1995, S. 120 f.

⁹⁵ vgl. ebenda

statt zu beschneiden, wenn mit ihr neue Handlungsalternativen und Problemlösestrategien generiert werden können. Bei der Ablehnung von Bewerbern im Zuge der Personalauswahl und –entwicklung ist die Einschränkung des Handlungsspektrums nicht auf den Einsatz der psychologischen Instrumente zurückzuführen, sondern darauf, dass eine Entscheidung über beschränkte Ressourcen zu treffen ist.

Die genannten Grundsätze sollten von den Unternehmen in jedem Fall eingehalten werden, um juristisch und ethisch unbedenkliche Berufseignungsdiagnostik durchführen zu können. Die Diskussionen ob und wie Psychodiagnostik im Wirtschaftskontext durchgeführt werden soll, werden dennoch weitergehen. Angeheizt werden diese durch Unternehmen, die ihre Bewerberauswahl oder Personalentwicklung eher als Nebensache betrachten und diese ungeschulten Mitarbeitern überlassen.

2.7.3 Durchführung testgestützter Eignungsuntersuchungen

Die Durchführung von Eignungsuntersuchungen sollte auf jeden Fall von geschultem Personal übernommen werden. Grundsätzlich ist die Betreuung durch einen Psychologen ratsam, der mit entsprechender diagnostischer Ausbildung in der Lage ist, die Testergebnisse ordentlich auszuwerten und zu interpretieren.

Bei der Auswahl und Zusammenstellung von Testverfahren bezüglich einer bestimmten diagnostischen Fragestellung sind u.a. die folgenden Aspekte zu beachten:

- Welche Konstrukte soll das Instrument erfassen?
- Welches Verfahren ist von seiner psychometrischen Qualität und sozialen Akzeptanz geeignet, um zur Beantwortung der jeweiligen diagnostischen Fragestellung *sinnvoll* beizutragen?⁹⁶

Im deutschsprachigen Raum existieren leider nur wenige persönlichkeitsbeschreibende Verfahren, die durch gute wissenschaftliche Fundierung und einem engen Bezug zu den praktischen Aufgabenstellungen der

⁹⁶ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 49 f.

Eignungsdiagnostik gekennzeichnet sind. Viele Persönlichkeitsverfahren sind aus der Grundlagenforschung entstanden und die Konzepte oftmals überholt. Weiterhin weisen diese Verfahren Überschneidungen mit klinischen Ansätzen auf. Folglich ist der Einsatz im wirtschaftlichen Kontext häufig aus rechtlichen und ethischen Gründen bedenklich. Trotzdem gibt es einige Verfahren, die für die diagnostische Arbeit in der Wirtschaft in Betracht gezogen werden können.⁹⁷ In diesem Zusammenhang werden vor allem das Leistungsmotivations-Inventar (LMI, 2001) von Schuler und Prochaska und das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP, 2003) von Hossiep und Paschen häufig genannt. Beide Verfahren zeichnen sich u.a. durch die berufsbezogene Formulierung der Items und durch die Beschreibung der Persönlichkeit in wichtigen berufsrelevanten Bereichen aus. Durch eben diese berufsbezogene Erfassung der Persönlichkeit sind diese Verfahren im wirtschaftlichen Kontext als rechtlich und ethisch unbedenklich einzuordnen. Sie orientieren sich aber explizit an lebensälteren Personen und die Testbearbeitung setzt teilweise Berufserfahrung voraus. Für die Berufseignungsdiagnostik mit Jugendlichen und jungen Erwachsenen fehlen entsprechende Verfahren. Ein neues Verfahren, der Start-P (2010) von Beauducel und Kersting soll genau diese Lücke schließen. In diesem Verfahren werden situative Komponenten, die für Jugendliche und junge Erwachsene relevant sind (z.B. Schulsituationen, Lehre) verwendet, um allgemeine Verhaltenstendenzen zu erfassen. Diese sollen für die Qualität des beruflichen Verhaltens einen prognostischen Wert haben. Auch dieses Verfahren ist durch seine Berufsbezogenheit als rechtlich und ethisch unbedenklich einzustufen.⁹⁸

Die meisten psychologischen Testverfahren – insbesondere Persönlichkeitsverfahren – eignen sich für die Einzel- und für die Gruppendurchführung. Gerade das Wort „Test“ wird meist mit einer Prüfungssituation assoziiert und sollte vermieden werden. Man sollte eher darauf hinweisen, dass sich alle Fragen auf persönliche Einstellungen und Verhaltensweisen der Teilnehmer beziehen und es damit keine richtigen oder falschen Antworten gibt. Der Diagnostiker sollte versuchen, eine vertrauensvolle und motivierende Atmosphäre zu schaffen. Den Testkandidaten sollte die Rolle des durchzuführenden Verfahrens im Kontext der gesamten Eignungsbeurteilung erklärt werden und darauf hingewiesen werden, dass auf der Grundlage der Ergebnisse dieses einzelnen Verfahrens keine Entscheidung getroffen werden wird. Weitere Hinweise zu einer fairen Testgestaltung finden sich z.B. bei Wottawa (1980).⁹⁹

⁹⁷ vgl. ebenda

⁹⁸ vgl. <http://www.kersting-internet.de/start-p.html>, verfügbar am 19.08.2010

⁹⁹ vgl. ebenda

Jede Bewerbungssituation stellt für die Kandidaten eine Belastung da. Eine eignungsdiagnostische Situation ist dabei unter anderem durch folgende Faktoren gekennzeichnet:

➤ *Asymetrische Selbstenthüllung:*

Es findet kein wechselseitiger, offener Austausch mit den Kandidaten statt. Es werden nur einseitig Informationen über den Kandidaten gesammelt.

➤ *Vertrauen auf Vertraulichkeit:*

Nur wenn der Kandidat sicher sein kann, dass seine Daten absolut vertraulich behandelt werden, wird er in einem Persönlichkeitsverfahren auch sozial unerwünschte Verhaltensweisen einräumen.¹⁰⁰

➤ *Intransparenz des Urteils:*

Vor allem bei Persönlichkeitsverfahren ist für den Laien kaum nachvollziehbar, wie die Antworten auf die einzelnen Fragen zum Gesamtergebnis zusammengefasst werden.

Die genannten Aspekte lassen die Bewerber die eignungsdiagnostische Situation eher als unerfreulich erleben. Wie sie die Auswahlverfahren beurteilen hängt zudem davon ab, wie das gesamte Auswahlverfahren für sie verläuft. Da negative Bescheide statistisch um ein Vielfaches häufiger als Vertragsangebote sind, werden Auswahlverfahren in der öffentlichen Meinung eher negativ beurteilt.¹⁰¹ Negative Urteile oder Vorurteile können durchaus zu einem Problem für das Unternehmen werden, wenn es dadurch zu einer ungünstigen Selbstselektion von potentiellen Bewerbern kommt. Aus diesem Grund sollte jedes Auswahlverfahren auch hinsichtlich der Akzeptanz der eignungsdiagnostischen Situation bei den Bewerbern überprüft werden.

2.8 Motivationale Einflüsse und Fehlerfaktoren bei der Bearbeitung von Persönlichkeitsverfahren

In der Forschung konnten Faktoren identifiziert werden, die für die Bearbeitung von Testfragen durchaus Relevanz haben. Im Folgenden werden Faktoren aufgezeigt, die

¹⁰⁰ vgl. Anastasi, 1985, S. 1 ff.

¹⁰¹ vgl. Fruhner/ Schuler/Funke; 1991, S. 170 ff.

häufig diskutiert und mitunter als problematisch für den praktischen Einsatz deklariert werden.

2.8.1 Beantwortungen im Sinne sozialer Erwünschtheit

Der Aspekt der Sozialen Erwünschtheit und die daraus folgenden Effekte werden häufig als größtes Problem von Persönlichkeitsverfahren in Auswahl-situationen angesehen. Ist ein Testverfahren relativ transparent, können auch Laien relativ schnell erkennen, was die jeweilige Frage erfassen will. Man spricht auch von Augenscheinvalidität der Items. Dies eröffnet dem Kandidaten die Möglichkeit, sich so darzustellen, wie er in der jeweiligen Situation gern gesehen werden würde. Die Möglichkeit die Verfahren in dieser Weise zu beeinflussen, hat einige Autoren zu der Auffassung gelangen lassen, dass Persönlichkeitsverfahren grundsätzlich für jeden Selektionsprozess ungeeignet sind, so z.B. Hampel & Klinkhammer, 1978.¹⁰²

Zur Erforschung der sozialen Erwünschtheit wurden unterschiedliche Untersuchungen vorgenommen. Die Ergebnisse dieser Untersuchungen sind eindeutig und nicht überraschend: Erhalten Menschen eine entsprechende Instruktion, können sie ihre Antworten in die gewünschte Richtung manipulieren.¹⁰³ Weiterhin werden verschiedene Strategien vorgeschlagen, um das Problem der sozialen Erwünschtheit zu kontrollieren:

- Es werden Forced-Choice-Items entwickelt. Hier weisen die Wahlalternativen gleiche soziale Erwünschtheit auf.¹⁰⁴
- Damit die Testkandidaten nicht unmittelbar auf das erfragte dahinterliegende Konstrukt schließen können, werden die Items möglichst subtil formuliert. Die Items werden ohne Rückgriff auf ein bestimmtes Konstrukt entwickelt und sollen lediglich in der Lage sein, zwischen sich empirisch ergebenden verschiedenen Gruppen zu differenzieren. Dieser Ansatz ist jedoch auch mit einigen Problemen verbunden und es existieren kaum überzeugende Belege, dass diese Items tatsächlich weniger anfällig für soziale Erwünschtheit sind. Ein Zusammenhang zu beruflichen Anforderungen kann man kaum herstellen.

¹⁰² vgl. Hampel/ Klinkhammer, 1978, S. 58 ff.

¹⁰³ Hough, Eaton, Dunette, Kamp & McCloy, 1990 und Thornton & Gierasch, 1980

¹⁰⁴ vgl. Paulhus, 1986, S. 153

- Den Testteilnehmern wird mitgeteilt, dass Falschantworten als solche identifiziert werden und vor den Konsequenzen nicht wahrheitsgemäßer Beantwortung gewarnt. Die Wirkung dieser Strategie wurde durch Studien nachgewiesen. Dieses Vorgehen beinhaltet allerdings eine Täuschung und ist damit in Auswahl-situationen als ethisch bedenklich einzustufen.
- Eine weitere, nicht unbestrittene Methode, ist das Beifügen von speziellen Skalen, die soziale Erwünschtheit oder Lügentendenzen erfassen sollen. Es gibt einige Persönlichkeitsverfahren, die solche Skalen beinhalten. Teilweise existieren Verrechnungsregeln, wie die Skalenwerte bei entsprechenden Ausprägungen der sozialen Erwünschtheit zu korrigieren sind – nach Cronbach funktioniert das nicht: „If the subject tells lies to the tester there is no way to convert the lies to truth“.¹⁰⁵

Es existieren noch weitere Strategien, die aus Komplexitätsgründen hier nicht genannt werden, jedoch bei Hossiep, Paschen und Mühlhaus (2000) nachgelesen werden können.

Obwohl ausreichende Belege dafür vorliegen, dass Menschen ihre Testergebnisse bei entsprechender Instruktion hinsichtlich sozialer Erwünschtheit verändern können, ist jedoch entscheidend, ob sie das in den realen Auswahl-situationen auch wirklich tun. In einer Studie von Hough et al. (1990) wurde nachgewiesen, dass das Ausmaß der sozialen Erwünschtheit sehr gering ist. Hampel und Klinkhammer (1978) können diese Ergebnisse in ihren Studien replizieren. In „Ernstsituationen“ gewinnen möglicherweise Faktoren einen Einfluss, die einer Verfälschungsneigung entgegenwirken (z.B. die Befürchtung, dass Verfälschungen durchschaut werden). Das Auftreten sozialer Erwünschtheit ist jedoch auch in Auswahl-situationen nicht auszuschließen, es ist lediglich fraglich, von welcher Tragweite und Relevanz der Effekt wirklich ist¹⁰⁶

Weiterhin wurde in mehreren Untersuchungen nachgewiesen, dass die Kontrolle sozialer Erwünschtheit keinen positiven Effekt auf die Validität hat. Die positive Selbstdarstellung in Bewerbersituationen ist nicht zwangsläufig eine Form von Unsicherheit, Lügen oder absichtlicher Täuschung, sondern die beste Möglichkeit an die gegebenen Umstände zu adaptieren. Die zentrale Anforderung in Bewerbungssituationen ist sich möglichst positiv darzustellen – soziale Erwünschtheit ist damit kein Phänomen, das ausschließlich bei der Bearbeitung von

¹⁰⁵ Cronbach, 1970; S. 510

¹⁰⁶ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 62 f.

Persönlichkeitsverfahren wirksam wird. Es spricht für eine intelligente Adaption an die Erfordernisse der jeweiligen Situation, sich in bestimmten Kontexten sozial erwünscht zu verhalten. Eine Überstrapazierung dieses Verhaltens dagegen wird eher Skepsis hervorrufen.¹⁰⁷

2.8.2 Zustimmungstendenzen

Zustimmungstendenzen sind neben der sozialen Erwünschtheit ebenfalls untersucht worden. Man spricht auch von der Ja-Sage-Tendenz. Damit ist gemeint, dass ein Item weitgehend unabhängig von seinem Inhalt eher bejaht denn verneint wird. Zustimmungstendenzen sind wahrscheinlicher, wenn die Testkandidaten dem Iteminhalt indifferent gegenüberstehen. Es ist außerdem maßgebend, ob die Items zum Erfahrungsbereich der Probanden gehören, also überhaupt sinnvoll beantwortbar sind. Zusammenfassend kann man sagen, dass Zustimmungstendenzen kaum Gefahren für die praktische diagnostische Arbeit bergen.

2.8.3 Bevorzugung extremer oder mittlerer Antwortkategorien

Die Bevorzugung extremer oder bestimmter Antwortkategorien stellt ein drittes Antwortmuster da. Es tritt natürlich ausschließlich bei Verfahren mit mehr als zwei Beantwortungsalternativen auf. Nach Bem und Allen (1974) unterscheiden sich Personen dahingehend, wie geeignet Persönlichkeitseigenschaften zu ihrer Beschreibung sind, d.h. wie stark situationsvariabel ihr Verhalten ist. Diejenigen Menschen, bei denen das Verhalten stärker situationsabhängig ist (es variiert von Situation zu Situation) als bei anderen, sind möglicherweise dadurch gekennzeichnet, dass sie eher mittlere Antwortkategorien bevorzugen. Im Gegensatz dazu, können Personen, die durch eine gewisse Rigidität (Starrheit) in ihrem Verhalten gekennzeichnet und damit weniger unabhängig von situationalen Einflüssen sind, zu den extremen Antwortkategorien neigen.¹⁰⁸ Wird diese Annahme bestätigt, wäre eine entsprechende Antworttendenz bereits eine wichtige diagnostische Situation.

Eine sorgfältige Itemformulierung in Form von eindeutigen und spezifischen Items stellt bisweilen den besten Schutz vor allen Response-Sets dar.¹⁰⁹

¹⁰⁷ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 63 f.

¹⁰⁸ vgl. Amelang/Borkenau, 1984, S. 96

¹⁰⁹ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 65

Im folgenden Abschnitt der Arbeit werden drei etablierte Persönlichkeitsverfahren vorgestellt. Zwei Verfahren, das NEO-Fünf-Faktoren-Inventar und der Eysenck Personality Questionnaire in der revidierten Kurzform, repräsentieren dabei den trade-orientierten Ansatz, welcher besonders von den Theoretikern vertreten wird. Dagegen steht der Myers-Briggs Typenindikator, welcher den Typologien zugeordnet werden kann. Dieses Verfahren ist durch seine einfache Handhabung und die guten Auswertungsmöglichkeiten sehr praxisnah und wird gern im Bereich der Personalentwicklung und im Teamtraining von Unternehmen und Unternehmensberatungen eingesetzt. Die meisten Wissenschaftler stehen dem Myers-Briggs Typenindikator und allgemein den Typologien allerdings sehr kritisch gegenüber.

3 Theoretische Darstellung der ausgewählten Testverfahren

3.1 Das NEO-Fünf-Faktoren Inventar (NEO-FFI) nach Costa und McCrae

3.1.1 Grundkonzept

Das NEO - Fünf-Faktoren-Inventar ist ein faktorenanalytisch konstruiertes Fragebogenverfahren und dient der Erfassung individueller Merkmalsausprägungen in den Bereichen Neurotizismus, Extraversion, Offenheit für Erfahrung, Verträglichkeit und Gewissenhaftigkeit. Er ist ein vergleichsweise kurzer Persönlichkeitsfragebogen mit 60 Items, welche auf einer fünfstufigen Skala von „starker Ablehnung“ bis „starker Zustimmung“ beantwortet werden. Diese 60 Items werden fünf Dimensionen mit je 12 Items zugeordnet.¹¹⁰

Das vorliegende Inventar ist eine deutsche Übersetzung des „NEO Five-Factor Inventory“ (NEO-FFI) von Paul Costa und Robert McCrae (1989,1992b). Es wurde 1993 von Peter Borkenau und Fritz Ostendorf übersetzt. Die Originalversion konnte aufgrund möglicher kultureller und sprachlicher Unterschiede nicht ungeprüft auf die deutsche Version übertragen werden. Deshalb schlossen sich an die Übersetzung zahlreiche Datenerhebungen an, mit denen eine hohe transkulturelle Generalisierbarkeit, der für die Originalversion berichteten Befunde, gezeigt werden konnte.¹¹¹

Durch das hohe Abstraktionsniveau ist das Inventar ein umfassendes, aber weniger spezifisches Instrument und eignet sich als Breitband- oder Screening- Verfahren zur Erfassung der Persönlichkeitsstruktur.¹¹² Es eröffnet einen groben, aber vollständigen Überblick über die Ausprägung der Persönlichkeit der Probanden auf den wichtigsten Dimensionen individueller Persönlichkeitsunterschiede und ist daher ein sehr nützliches Instrument für Forschungszwecke. Aufgrund seiner Bandbreite ist der Einsatz weiterhin im Bereich der klinischen Psychologie, der Berufsberatung und Organisationspsychologie und der Schullaufbahn- bzw. Studienberatung ebenfalls sinnvoll und denkbar.¹¹³

¹¹⁰ vgl. Borkenau/ Ostendorf, 1993; S. 5

¹¹¹ vgl. ebenda, S. 11

¹¹² vgl. Ziegler, 2002; S. 386

¹¹³ vgl. Borkenau/ Ostendorf; 1993; S. 22-23

3.1.2 Theoretische Begründung des Verfahrens

Grundlage des NEO-FFI ist der psycho-lexikalische Ansatz. Dieser lässt sich aus der Sedimentationshypothese von Ludwig Klages (1926) und Raymond B. Cattell (1943) ableiten.

Der Gedanke des psycho-lexikalischen Ansatzes ist, dass alle Persönlichkeitsmerkmale ihr Abbild in der Sprache finden. Nachdem man zunächst möglichst viele Adjektive zur Beurteilung der menschlichen Persönlichkeit gesammelt hat, versucht man in einem zweiten Schritt, die Gesamtheit dieser Adjektive auf empirischem Weg zu Faktoren zu verdichten. Diese Faktoren bilden die elementaren Persönlichkeitsdimensionen ab.¹¹⁴ „Zahlreiche, in der Forschungstradition Cattells stehende Autoren haben nach den sprachlich repräsentierten Eigenschaften gesucht, mit denen sich die Unterschiede zwischen Personen hinreichend genau beschreiben lassen.“¹¹⁵ Während Cattell zunächst zwei Faktoren identifizierte und seinen Fragebogen später noch um vier weitere Skalen ergänzte, wählte Eysenck ein Modell mit nur drei Faktoren. Im Zuge einer Realanalyse der Daten Cattells wurde schon frühzeitig eine Lösung mit fünf Faktoren vorgeschlagen. Diese stellte sich auch in anderen, teilweise früheren, faktorenanalytischen Verfahren als angemessen heraus und wurde später von zahlreichen Befunden untermauert.¹¹⁶

Als Ergebnis dieser verschiedenen Datenreduktionen kristallisierten sich demnach fünf Faktoren heraus, die sogenannten „Big Five“, die die Grunddimensionen individueller Unterschiede darstellen.¹¹⁷

Aus der Entwicklung der amerikanischen NEO-Inventare gingen insgesamt drei Versionen des NEO hervor: das NEO-Persönlichkeitsinventar (PI; Costa & McCrae, 1985), das NEO-Fünf-Faktoren-Inventar (NEO-FFI; Costa & McCrae, 1992) und das revidierte NEO-Persönlichkeitsinventar (NEO-PIR; Costa & McCrae, 1992; Costa, McCrae & Dye, 1991).¹¹⁸

¹¹⁴ vgl. Ziegler, 2002, S. 386

¹¹⁵ Hossiep/ Paschen/ Mühlhaus, 2000, S. 117

¹¹⁶ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 117

¹¹⁷ vgl. Ziegler, 2002, S. 386

¹¹⁸ vgl. ebenda

3.1.3 Die Entwicklung des Verfahrens

1986 begannen Borkenau und Ostendorf mit der Übersetzung des Persönlichkeitsinventares, zunächst mit dem NEO-PI, um auch dem deutschsprachigen Raum ein Verfahren zugänglich zu machen, das die Dimensionen des Big-Five-Modells umfasst. Die Zugehörigkeit der Items zu den fünf Faktoren und die Bedeutung dieser, sollte durch die Übersetzung in die deutsche Sprache nicht verändert werden. Daher wurde sinngemäß und nicht wortgetreu übersetzt. Alle 60 Items des NEO-FFI sind ebenfalls im NEO-PI enthalten. Für die Entwicklung der deutschen Version des NEO-FFI mussten diese Items lediglich gesondert analysiert werden. Die gesonderte Übersetzung des amerikanischen NEO-FFI bzw. eine neue Datenerhebung sahen die Autoren als nicht erforderlich an. In mehreren Untersuchungen wurde festgestellt, dass die Korrespondenz zur amerikanischen Version vollständig erhalten blieb.¹¹⁹

3.1.4 Skalen und Interpretationshinweise

In der folgenden Tabelle sind die Dimensionen des NEO-FFI dargestellt. Die Skalencharakterisierungen wurden dem Testmanual entnommen. Jede Skala umfasst zwölf Items. Die Zuordnung der Items zu den Skalen und ausführlichere Interpretationshinweise kann man ebenfalls im Manual nachlesen.

Tab. 2: Die fünf Merkmalsbereiche des NEO-FFI (nach Borkenau & Ostendorf, NEO-Fünf-Faktoren-Inventar nach Costa und McCrae)

Merkmalsbereich	Beschreibung
Neurotizismus	Probanden mit hohen Werten in <i>Neurotizismus</i> neigen dazu, nervös, ängstlich, traurig, unsicher und verlegen zu sein und sich Sorgen um ihre Gesundheit zu machen. Sie neigen zu unrealistischen Ideen und sind weniger in der Lage, ihre Bedürfnisse zu kontrollieren und auf Stresssituationen angemessen zu reagieren.
Extraversion	Probanden mit hohen Werten in <i>Extraversion</i> sind gesellig, aktiv, gesprächig, personenorientiert, herzlich, optimistisch und heiter. Sie mögen Anregungen und Aufregungen.

¹¹⁹vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 119

Offenheit für Erfahrung	Probanden mit hohen Werten in <i>Offenheit für Erfahrung</i> zeichnen sich durch eine hohe Wertschätzung für neue Erfahrungen aus, bevorzugen Abwechslung, sind wissbegierig, kreativ, phantasievoll und unabhängig in ihrem Urteil. Sie haben vielfältige kulturelle Interessen und interessieren sich für öffentliche Ereignisse.
Verträglichkeit	Probanden mit hohen Werten in der Skala <i>Verträglichkeit</i> sind altruistisch, mitfühlend, verständnisvoll und wohlwollend. Sie neigen zu zwischenmenschlichem Vertrauen, zur Kooperativität, zur Nachgiebigkeit, und sie haben ein starkes Harmoniebedürfnis.
Gewissenhaftigkeit	Die Skala <i>Gewissenhaftigkeit</i> schließlich unterscheidet ordentliche, zuverlässige, hart arbeitende, disziplinierte, pünktliche, penible, ehrgeizige und systematische von nachlässigen und gleichgültigen Personen.

3.1.5 Normen und Gütekriterien

Da bisher noch keine repräsentative Stichprobe vorliegt, wurden von den Autoren noch keine Normen veröffentlicht. Aus dem gleichen Grund wurden auch keine Standardwerte oder Prozentränge angegeben.¹²⁰

Objektivität

Die von Beate Ziegler beschriebenen Probleme hinsichtlich der Durchführungsobjektivität können von der Autorin dieser Arbeit nicht bestätigt werden. Ziegler sieht die Durchführungsobjektivität mangels präziser Angaben gefährdet. Aufgrund des ausführlichen Manuals ist der standardisierte Ablauf der Durchführung gewährleistet. Im empirischen Teil der Arbeit wurde der NEO-FFI außerdem in der computerbasierten Form durchgeführt, so dass hinsichtlich der Durchführungsobjektivität keinerlei Bedenken bestehen.

Die Auswertungsobjektivität ist bei der Papierform aufgrund einer standardisierten Auswertung per Schablone, genauso wie beim computerbasierten Verfahren, ebenfalls gegeben.

¹²⁰ vgl. Borkenau/ Ostendorf, 1993, S. 22

Zur Interpretation individueller Messwerte liegt eine ausführliche verbale Skalenbeschreibung und Mittelwertsangaben vor. Auch die Interpretationsobjektivität ist somit eingeschränkt gewährleistet. In der computerbasierten Version wurde diese gesteigert, indem konkrete Ergebnisprofile dargestellt werden.¹²¹

Reliabilität

Die Reliabilität des NEO-FFI wird mittels Cronbach's alpha ausgewiesen. Dabei sind die hohen internen Konsistenzen der Skalen von .71 bis .85 unter der Berücksichtigung der geringen Itemanzahl je Skala beachtlich.

Die Berechnung der Retest-Reliabilität erfolgte auf der Basis einer Teilstichprobe von 146 Personen, die den Fragebogen nach Ablauf von zwei Jahren erneut beantworteten. Die Retest-Reliabilität wird zwischen .65 und .81 angegeben.¹²²

Die folgende Tabelle gibt nochmals einen Überblick über die Reliabilitäten des NEO-FFI.

Tab. 3: Die Reliabilitäten des NEO-FFI (vgl. Borkenau & Ostendorf, 1993)

Skala	Cronbachs α (N = 2112)	Retest-Korrelation nach 2 Jahren
Neurotizismus	.85	.80
Extraversion	.80	.81
Offenheit für Erfahrung	.71	.76
Verträglichkeit	.71	.65
Gewissenhaftigkeit	.85	.81

Validität

Als Beleg für die Konstruktvalidität liegen neben den Testwerten auch Selbsteinschätzungen und Bekanntenratings von 300 Teilnehmern vor. Die Einschätzungen erfolgten auf siebenstufigen Adjektivskalen, wobei die Bekannten ihre Urteile unter Beaufsichtigung eines Versuchsleiters abgaben, um die experimentelle Unabhängigkeit von Selbst- und Bekanntenbeurteilungen zu gewährleisten. Die Korrelationen der Werte des NEO-FFI mit den Bekanntenratings streuen zwischen .27

¹²¹ vgl. Ziegler, 2002, S. 389

¹²² vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S.120

und .45. Diese Validitätskoeffizienten dürften durch den Umstand gemindert sein, dass nicht immer die nächsten Bekannten und Verwandten in die Universität mitgebracht werden konnten, sodass unter Umständen weniger nahestehende Personen die Bekannteneinschätzungen abgaben. Damit dürfte es sich bei diesen Werten um eher zu niedrige Validitätsschätzungen handeln. Bei einer postalischen Befragung näherer Bekannter von 116 Teilnehmern lagen die konvergenten Validitäten auf einem deutlich höheren Niveau. Diese Koeffizienten streuten zwischen .43 und .61. Bei dieser Befragung konnte jedoch die Unabhängigkeit von Selbst- und Fremdeinschätzung - aufgrund der postalischen Befragung - nicht sichergestellt werden, sodass die Validitäten unter Umständen überschätzt wurden.¹²³

Aufgrund des Konstruktionsprinzips des NEO-FFI sehen die Autoren keine Erfordernis, Studien zur Kriteriumsvalidität vorzulegen. Die Aufgabe des NEO-FFI sehen sie vielmehr in der Erfassung faktorenanalytisch gewonnener Konstrukte und nicht in der Vorhersage spezifischer Kriterien.¹²⁴

3.2 Der Myers-Briggs Typenindikator (MBTI)

3.2.1 Grundkonzept

Der MBTI gehört ebenso wie der NEO-FFI zu den Persönlichkeits-Struktur-Tests. Er ist ein Selbstbeurteilungsverfahren, das auf der Basis der Typenlehre von Carl Gustav Jung entwickelt wurde, um die Theorie für die Praxis nutzbar zu machen.¹²⁵

Der MBTI ist in drei Teile untergliedert und umfasst insgesamt 88 Items. Der erste und dritte Teil bestehen aus Fragen mit Alternativantworten. Im zweiten Teil muss man je Item aus zwei vorgegebenen Wörtern eines auswählen. Der MBTI besteht aus acht Skalen.

¹²³ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S.120

¹²⁴ vgl. ebenda

¹²⁵ vgl. Deubner, 2002, S. 692 f.

Zwei Skalen erfassen dabei jeweils die zwei gegensätzlichen Pole einer persönlichen Präferenz:

S	„Sensing“	Sinnliche Wahrnehmung
N	„Intuition“	Intuitive Wahrnehmung
T	„Thinking“	Analytische Beurteilung
F	„Feeling“	Gefühlsmäßige Beurteilung
E	„Extraversion“	
I	„Introversion“	
J	„Judging“	Beurteilung
P	„Perceiving“	Wahrnehmung

Die verschiedenen Skalen beschreiben Dichotomien. Damit sind 16 verschiedene Persönlichkeitstypen erfassbar. Eine Dominanz kann auftreten hinsichtlich der sinnlichen versus der intuitiven Wahrnehmung, der analytischen versus der gefühlsmäßigen Beurteilung, der Extraversion versus der Introversion und der Beurteilung versus der Wahrnehmung. Je nach Präferenz der einen oder anderen Skala kann man spezifische dynamische Beziehungen feststellen, anhand derer sich jeweils der Persönlichkeitstyp bestimmen lässt.¹²⁶

Der MBTI kann vielfältig eingesetzt werden, findet jedoch in der Beratung seine Hauptanwendung. Er kann beispielsweise bei Schulungen von Führungskräften, Kommunikationstrainings, der Personalberatung oder der Diagnostik und Verbesserung von Teamarbeit eingesetzt werden.

3.2.2 Theoretischer Hintergrund

Die Typentheorie des Schweizer Arztes und Psychoanalytikers Carl Gustav Jung dient dem MBTI als Ausgangspunkt und Grundlage. Jung entwickelte seine Theorie zur Erklärung der normalen Unterschiede zwischen gesunden Menschen. Nach dieser Theorie kann man das Verhalten eines Menschen durch seine individuellen

¹²⁶ vgl. Deubner, 2002. S. 692 ff.

Präferenzen, seine spezifische Herangehensweise und Interpretation der Wirklichkeit verdeutlichen.¹²⁷

Die Jungsche Theorie des psychologischen Typus konstatiert acht verschiedene Muster normalen Verhaltens. Diese werden als Typen bezeichnet. Nach Jung ist der gesunde Mensch, wenn sein Verstand und seine Sinne aktiv sind, mit einer von den zwei typischen Tätigkeiten beschäftigt, nämlich mit der Aufnahme von Informationen (wahrnehmen) oder mit dem Ordnen dieser Informationen und entsprechenden Schlussfolgerungen (urteilen). Jung stellte fest, dass der Mensch auf zwei gegensätzliche Arten wahrnimmt, die er als *Empfindung* und *Intuition* bezeichnete und dass er auf zwei gegensätzliche Arten urteilt, die er *Denken* und *Fühlen* nannte.¹²⁸

Jung beobachtete weiterhin, dass ein Mensch seine Energie entweder mehr auf die Außenwelt der Menschen richtet (Erfahrungen und Tätigkeiten) und von dieser angeregt wird, oder dass er sie mehr auf die Innenwelt (Gedanken, Erinnerungen, Emotionen) richtet und durch diese mehr angeregt wird. Er benannte diese zwei Richtungen der Energie mit *Extraversion* (Handeln in der Außenwelt) und *Introversion* (Reflektieren in der Innenwelt).¹²⁹

Jeder der vier beschriebenen mentalen Prozesse - Empfinden, Intuition, Denken, Fühlen - besitzt seine eigenen prädiktiven Eigenschaften, prägt sich jedoch unterschiedlich aus, je nachdem ob er mehr auf die extravertierte oder die introvertierte Welt gerichtet ist. Durch die Kombination der vier mentalen Prozesse und der zwei unterschiedlichen Ausrichtungen gegenüber der Welt charakterisiert Jung acht basale Muster psychischer Tätigkeit, die dem Menschen zur Verfügung stehen.¹³⁰

Diese acht psychischen Prozesse stehen jedem Menschen zur Verfügung und werden von jedem Menschen benutzt. Jedoch bestehen angeborene Unterschiede, was die Bevorzugung der jeweiligen Prozesse betrifft. Die natürliche Präferenz für eine dieser Funktionen gegenüber den anderen, gibt Menschen vor, ihre Energie auf eben diese Funktion zu richten. So entwickeln sie Verhaltensweisen und Persönlichkeitsmuster, die für diese Funktion charakteristisch sind. Jung bezeichnete den präferierten psychischen Prozess als die dominante Funktion.¹³¹

¹²⁷ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 124 f.

¹²⁸ vgl. Kirby/ Myers, 2001, S. 6

¹²⁹ vgl. ebenda,

¹³⁰ vgl. ebenda, S. 7

¹³¹ vgl. ebenda

Die Unterschiede in der bevorzugten, benutzten und entwickelten psychischen Funktion führen zu basalen Unterschieden zwischen den Menschen. Es ergeben sich vorhersagbare Verhaltensmuster, welche in einem dynamischen Wechselspiel mit den anderen psychischen Funktionen die psychologischen Typen ergeben. Jung konzentrierte sich bei seiner Arbeit auf die genaue Beschreibung der acht dominanten Funktionen. Er erkannte weiterhin, dass die Menschen die Funktionen in einer Art Hierarchie benutzten. Jeder individuelle Typ benutzt sie dabei in einer bestimmten Reihenfolge. Jung nannte die Funktionen in folgender Ordnung: die dominante Funktion - der erste, am meisten benutzte psychische Prozess; die Hilfsfunktion, der zweite bevorzugte Prozess; die tertiäre und die inferiore Funktion schließen sich als am wenigsten bevorzugte Funktionen an.¹³²

1926 veröffentlichte Katherine Briggs unabhängig von Jung eine eigene Klassifikation der Persönlichkeit. Aber auch sie ging davon aus, dass Menschen sich vor allem durch ihre Zugänge zum Leben unterscheiden. Sie stieß eher zufällig auf frühere Publikationen Jungs, entdeckte die Parallelen ihrer beiden Theorien und beschäftigte sich nun intensiver mit seiner Konzeption. Bei ihren Forschungsarbeiten wurde sie von ihrer Tochter Isabel Myers-Briggs unterstützt. Der Ausbruch des Zweiten Weltkrieges und der Nationalsozialismus in Deutschland beeinflussten die Forschungsarbeiten der beiden Frauen maßgeblich. Sie kamen zu der Überzeugung, dass zahlreiche Konflikte maßgeblich aus der Unfähigkeit resultieren, die Unterschiede zwischen den Menschen zu verstehen und damit umgehen zu können. Ziel der Forschungsarbeiten war ein Fragebogen, mit dem man die interindividuellen Differenzen der Wahrnehmung und Beurteilung der Wirklichkeit messen kann. Nachdem sie 1941 mit der Item-Sammlung begannen, wurden in den 40er und 50er Jahren ausführliche Testungen gemacht. 1962 wurde schließlich die erste Forschungsversion des Myers-Briggs-Typenindikators vom „Education Testing Service“ veröffentlicht. Für den Einsatz in der Beratung und der praktischen Arbeit wurde das Verfahren erst 1975 - nach 30-jährigen Forschungsbemühen - als reif angesehen und publiziert. Isabel Myers-Briggs hat bis zu ihrem Tod 1982 an der Weiterentwicklung ihres Instrumentes gearbeitet.¹³³

Briggs und Myers entwickelten die Jungsche Vorstellung von der Hilfsfunktion weiter. Sie nahmen sie in ihr Konzept auf und konzipierten schließlich die vom MBTI angezeigten 16 Typen, die in der nachfolgenden Tabelle noch einmal übersichtlich dargestellt sind.¹³⁴

¹³² vgl. Kirby/ Myers, 2001, S. 6 ff.

¹³³ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 125

¹³⁴ vgl. Kirby/ Myers, 2001, S. 7

Tab. 4: Die 16 Typen des MBTI (nach Kirby/ Myers, 2000, S. 7)

Dominante Funktion	Hilfsfunktion	MBTI-Typ
Introvertiertes Empfinden	mit Extravertiertem Denken	ISTJ
Introvertiertes Empfinden	mit Extravertiertem Fühlen	ISFJ
Extravertiertes Empfinden	mit Introvertiertem Denken	ESTP
Extravertiertes Empfinden	mit Introvertiertem Fühlen	ESFP
Introvertierte Intuition	mit Extravertiertem Denken	INTJ
Introvertierte Intuition	mit Extravertiertem Fühlen	INFJ
Extravertierte Intuition	mit Introvertiertem Denken	ENTP
Extravertierte Intuition	mit Introvertiertem Fühlen	ENFP
Introvertiertes Denken	mit Extravertiertem Empfinden	ISTP
Introvertiertes Denken	mit Extravertierter Intuition	INTP
Extravertiertes Denken	mit Introvertiertem Empfinden	ESTJ
Extravertiertes Denken	mit Introvertierter Intuition	ENTJ
Introvertiertes Fühlen	mit Extravertiertem Empfinden	ISFP
Introvertiertes Fühlen	mit Extravertierter Intuition	INFP
Extravertiertes Fühlen	mit Introvertiertem Empfinden	ESFJ
Extravertiertes Fühlen	mit Introvertierter Funktion	ENFJ

3.2.3 Entwicklung des Verfahrens

Die Entwicklung des deutschen Verfahrens begann mit der Übersetzung der insgesamt 290 Items der englischen Version, bei der eine sprachlich-kulturelle Anpassung vorgenommen, aber ein möglichst großes Gleichmaß zu den Ursprungsitems angestrebt wurde. Die Testfragen sind im forced-choice-Format vorgegeben: Der Proband wählt zwischen zwei vorgegebenen Antwortalternativen, welche jeweils eine spezifische Präferenz repräsentieren. Aufgrund dessen war darauf zu achten, dass beide Antwortalternativen gleichwertig sind und keine Präferenz als minderwertig oder überlegen betrachtet wird. Das Verfahren sollte nicht im klinischen Bereich eingesetzt werden. Es kam also nicht darauf an, extrem seltene oder ungewöhnliche

Einstellungen oder Verhaltensweisen zu betrachten. Vielmehr kam es auf eine hohe Messgenauigkeit im „Durchschnittsbereich“ an. Hier ist die Unterscheidung der Präferenzen am schwierigsten.¹³⁵

Mit Hilfe von Itemanalysen wurden letztlich 88 aus ursprünglich 290 Items ausgewählt, die optimal zwischen den beiden Polen einer Präferenzskala differenzieren und gleichzeitig nur mit einer der vier Skalen korrelieren. Eine völlige Unabhängigkeit der Skalen konnte nicht erreicht werden. Damit verblieben gewisse Zusammenhänge zwischen den Dimensionen.¹³⁶

3.2.4 Skalen und Interpretationshinweise

Die Autoren haben detaillierte Konzepte für die Rückmeldung der Ergebnisse an die Probanden entwickelt, da der MBTI vor allem für die Beratung empfohlen wird. Das Testmanual gibt umfassende Hinweise für die Besprechung der Ergebnisse. Die ausführlichen Typenbeschreibungen im Manual sollen dem Teilnehmer vorgelegt werden und als Grundlage für das anschließende Gespräch dienen. Weiterhin sind zahlreiche Leitfäden von OPP erschienen, die die Konzepte der Typentheorie noch einmal aufgreifen und hinsichtlich verschiedener Themen interpretieren; z.B: „Einführung von Typen innerhalb von Organisationen“.

Für das Feedback sollten nicht die Dimensionen, sondern der sich ergebende Typ ganzheitlich betrachtet werden. Um zu verdeutlichen, wie die Interpretation eines solchen Typs aussieht, wird an dieser Stelle beispielhaft in Auszügen die Charakterisierung eines ISFP-Typen wiedergegeben, der seine Energie aus der seiner Innenwelt bezieht, realistische und faktische Information mag, seine Entscheidungen aufgrund seiner persönlichen Werte trifft und ein flexibles und anpassungsfähiges Leben mag.¹³⁷

„Menschen mit ISFJ-Präferenzen sind zuverlässig und rücksichtsvoll, fühlen sich den Menschen und Gruppen, mit denen sie in Verbindung stehen, verpflichtet, übernehmen Verantwortung. Sie arbeiten mit beständiger Energie daran, Aufgaben vollständig und rechtzeitig zu erfüllen. Sie scheuen keine Mühe, wenn es darum geht, etwas

¹³⁵ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S.129

¹³⁶ vgl. ebenda

¹³⁷ vgl. Kirby/ Myers, 2001, S. 11

fertigzustellen, das sie als notwendig ansehen, haben es aber nicht gern, wenn sie etwas tun sollen, worin sie keinen Sinn sehen.

ISFJ-Menschen konzentrieren sich auf das, was Menschen brauchen und wollen, und stellen deshalb Regeln auf und führen Verfahren ein, um sicherzustellen, dass diese Bedürfnisse und Wünsche erfüllt werden. Sie nehmen Rollen und Verantwortung ernst und wollen, dass andere dies auch tun. Familiäre Beziehungen und Familienverantwortung sind äußerst wichtig für ISFJ-Menschen; sie erfüllen ihre Rollen gewissenhaft und erwarten dasselbe auch von den anderen Familienmitgliedern. (...)

Manchmal wurden ISFJ-Menschen durch ihre persönlichen Umstände nicht in der Entwicklung und im Ausdruck ihrer Fühl- und Empfindungspräferenzen unterstützt. (...)

Unter hoher Stressbelastung können sich ISFJ-Menschen in „Katastrophenvisionen“ verlieren. Sie können diese dann zum Thema machen, ohne sich wie sonst üblich darum zu kümmern, wie sich dies auf die Menschen um sie herum auswirken wird.“¹³⁸

Da mit standardisierten Textbausteinen gearbeitet wird, haben die Beschreibungen oft horoskopähnlichen Charakter. Dennoch dienen sie als gute Grundlage für ein auswertendes Gespräch. Es kann gemeinsam mit dem Teilnehmer reflektiert werden, für welche der angeführten Charakterisierungen er konkrete Verhaltensbeispiele nennen kann und für welche nicht.¹³⁹

3.2.5 Normen und Gütekriterien

Zur Normierung sind gegenwärtig Mittelwerte und Standardabweichungen für verschiedene Altersgruppen angegeben, welche jedoch nunmehr die amerikanische Normstichprobe charakterisieren. Eine Datenbank deutschsprachiger Teilnehmer befindet sich derzeit in der Entwicklung.¹⁴⁰

Objektivität

Die Durchführung ist schriftlich festgelegt und der standardisierte Fragebogen, der 88 Items umfasst und den Probanden jeweils zwei Antwortmöglichkeiten vorgibt, gewährleistet eine hohe Durchführungsobjektivität.

¹³⁸ Kirby/ Myers, 2001, S.15

¹³⁹ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 131

¹⁴⁰ vgl. ebenda

Die Auswertungsobjektivität wird durch die Anwendung einer Schablone bzw. eines streng vorgegebenen Auswertungsverfahrens, das im Manual beschrieben wird, sichergestellt.

Die Rohwerte werden mit der Schablone ermittelt und nach der Auszählung mittels vorgegebener Multiplikatoren zu korrigierten Werten verrechnet. Damit wird die unterschiedliche Anzahl von Items pro Skala berücksichtigt. Mit Hilfe einer einfachen Rechenvorschrift kann für jede Skala der präferierte Pol und somit der Typ bestimmt werden.¹⁴¹

Hinsichtlich der Interpretationsobjektivität ist der MBTI eher als mangelhaft einzustufen. Trotz der im Manual vorgegebenen Interpretationshinweise und Charakterisierungen der Typen bleibt sehr viel Spielraum für die Ergebnisinterpretation. Durch standardisierte Textbausteine und globale Formulierungen kann das Ergebnis zu unterschiedlich ausgelegt werden. Weiterhin sind keine Vergleichsgruppen vorhanden. Die mehr oder weniger ausgeprägte Zugehörigkeit zu einem Typ kann nicht exakt bestimmt werden.

Reliabilität

Die interne Konsistenz der einzelnen Skalen (N=548) nach Cronbachs Alpha beträgt $r = .92$ für EI, $r = .88$ für SN, $r = .87$ für TF und $r = .88$ für JP und kann damit als zufriedenstellend eingestuft werden.¹⁴²

Für die Retest-Reliabilität halten die Autoren die Korrelation der Rohwerte für weniger bedeutsam. Eine hohe Bedeutung im Sinne der Zuverlässigkeit und Genauigkeit messen sie einer identischen Typenklassifizierung nach einem bestimmten Zeitintervall zu. Zur Bestimmung dieser Retest-Reliabilität wurde an einer Stichprobe von $N = 40$ nach einem Zeitraum von sechs Wochen bestimmt, wie hoch die Proportion der Fälle mit identischer Buchstabenkombination ist. Vier identische Buchstaben, d.h. vollständige Übereinstimmung trat dabei in 67,5% der Fälle auf, drei identische Buchstaben lagen bei 25% der Fälle und zwei identische Buchstaben bei 7,5% der Fälle vor.¹⁴³

Die Werte liegen deutlich über den in den USA gewonnenen Kennwerten, was durchaus auf die geringe Stichprobengröße zurückzuführen ist.

¹⁴¹ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 132

¹⁴² vgl. Deubner, 2002, S. 694

¹⁴³ vgl. ebenda

Validität

Die Konstruktvalidität wurde mit Hilfe einer Typenübersicht überprüft, wobei, entsprechend der Theorie, davon ausgegangen wird, dass bestimmte Typen von Personen spezielle berufliche Interessen haben. Dahingehend wurden Fachhochschulstudenten der Betriebswirtschaft (N = 77), Restaurantleiter (N = 89) und Personalentwickler (N = 132) getestet. Es wurde deutlich, dass Anforderungsprofil, Tätigkeitsmerkmale und Typen in einer Beziehung zueinander stehen. Es kristallisierten sich Typenverteilungen heraus, die den Vorhersagen der Theorie entsprachen. Diese Ergebnisse werden als Hinweise auf die Validität bewertet, da die Typenbeschreibungen auch die Arbeitsweisen der Personen enthalten.¹⁴⁴

Um die Konstruktvalidität weitergehend abzusichern, wurde eine Vielzahl von Korrelationsbestimmungen der MBTI-Skalenwerte mit anderen Fragebögen (z.B. MMPI von Dahlstrom & Welsh; 16PF von Cattell et al.) bei verschiedenen Berufsgruppen durchgeführt. Hier zeigten sich einige substantielle Zusammenhänge, die in einer ausführlichen Tabelle in der Handanweisung (Seite 72 ff.) nachzulesen sind.¹⁴⁵

Die Pearson-Produktmoment-Korrelationen, welche über die acht Unterskalen bestimmt wurden, zeigen eine hohe Unabhängigkeit der Skalen an. Auch die angenommenen Negativkorrelationen zwischen einer Unterskala und ihrem Gegenpol wurden empirisch bestätigt. Eine Faktorenanalyse verifiziert die Annahme der Unabhängigkeit der einzelnen Skalen.¹⁴⁶

Alle erwähnten Ergebnisse beziehen sich auf angloamerikanische Untersuchungen. Es fanden weitere Untersuchungen statt, die die Theorie im Allgemeinen bestätigen.

3.3 Der Eysenck Personality Questionnaire

3.3.1 Grundkonzept

Der Eysenck Personality Questionnaire ist ein Instrument, das der Erfassung der Superfaktoren der Persönlichkeit nach Eysenck - Psychotizismus, Extraversion und

¹⁴⁴ vgl. Deubner, 2002, S. 694

¹⁴⁵ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 131

¹⁴⁶ vgl. Deubner, 2002, S. 694

Neurotizismus - dient. Zur Kontrolle der Antworttendenzen in Richtung sozialer Erwünschtheit dient eine vierte Skala. Das Persönlichkeitsmodell nach Eysenck bestimmt eine hierarchische Anordnung von Persönlichkeitseigenschaften, wobei die Superfaktoren (Persönlichkeitstypen oder Faktoren zweiter Ordnung) durch Eigenschaften (traits, Faktoren erster Ordnung) definiert werden.¹⁴⁷

Willibald Ruch konstruierte die vorliegende deutsche Kurzfassung des Eysenck Personality Questionnaire-revised (EPQ-RK). Diese enthält 50 Fragen, wobei Extraversion, Neurotizismus und die vierte Skala durch je 12 Fragen abgedeckt werden, während Psychotizismus durch 14 Fragen erfasst wird.

Da vor allem die Dimensionen Extraversion und Neurotizismus wichtig für die psychiatrische und medizinische Forschung sind, ist der EPQ-R bzw. der EPQ-RK vor allem im klinischen Bereich gut einsetzbar. Aber auch in der Markt- und Motivationsforschung und in Erziehungsberatungs- und Berufsberatungsstellen findet er seine Anwendung.

3.3.2 Theoretische Begründung des Verfahrens

Eysenck entwickelt ebenso wie Cattell ein hierarchisches Strukturmodell der Persönlichkeit, unterschied sich aber in seiner Vorgehensweise. Er war einerseits experimentell und naturwissenschaftlich orientierter ausgerichtet und wand andererseits die hypothetico-deduktive¹⁴⁸ Vorgehensweise an. Aufgrund seiner Forschungsarbeiten erklärte er viele Ausprägungsunterschiede in der Persönlichkeit und den daraus resultierenden Verhaltensunterschieden durch grundsätzliche Unterschiede in der Physiologie der Menschen.¹⁴⁹

„Persönlichkeit ist die mehr oder weniger feste und überdauernde Organisation des Verhaltens, des Temperaments, des Intellekts und der Physis eines Menschen und determiniert seine einzigartige Anpassung an die Umwelt.“¹⁵⁰

¹⁴⁷ vgl. Ruch, Die revidierte Fassung des Eysenck Personality Questionnaire und die Konstruktion des deutschen EPQ-R bzw. EPQ-RK, 1999

¹⁴⁸ Bei dem hypothetico-deduktiven Modell handelt es sich um ein einheitswissenschaftliches Modell, dessen Ausgangspunkt theoretische Überlegungen bzw. Annahmen sind, die an empirischen Daten überprüft werden.

¹⁴⁹ vgl. Ibrahimovic/ Bulheller, 2008, S. 13

¹⁵⁰ Ibrahimovic/ Bulheller, 2008, S.13

Eysenck nutzte, neben verschiedenen Informationsquellen zur Messung der Persönlichkeit, Lernexperimente (klassisches Konditionieren), um die Eigenschaften seines Strukturmodells zu verifizieren. Er unterscheidet in diesem Modell zwischen Traits und Typen. Traits werden dabei als Dispositionen für Erleben und Verhalten wie z.B. Geselligkeit oder Aktivität definiert. Typen dagegen sind kontinuierlich verteilte Persönlichkeitsdimensionen, welche eine biologische Basis haben und eine übergeordnete Kategorie in diesem Modell darstellen.¹⁵¹

Eysenck leitete aus den Ergebnissen von Persönlichkeitsverfahren drei breite Dimensionen von Eigenschaften (Typen) ab:

- *Extraversion*: Inwieweit ist eine Person nach innen oder nach außen orientiert?
- *Neurotizismus*: Inwieweit ist eine Person emotional stabil oder labil?
- *Psychotizismus*: Inwieweit ist eine Person freundlich und rücksichtsvoll oder aggressiv und antisozial?

Die ersten beiden Dichotomien lassen sich in einem Koordinatenkreuz, in dem theoretisch jeder Mensch untergebracht werden kann, darstellen. Jeder Quadrant entspricht dabei einem der vier Temperamente:

- Stabilität + Introversion = **Phlegmatiker**
- Stabilität + Extraversion = **Sanguiniker**
- Instabilität + Introversion = **Melancholiker**
- Instabilität + Extraversion = **Choleriker**

¹⁵¹ vgl. Ibrahimovic/ Bulheller, 2008, S.13

Die folgende Abbildung stellt das Modell Eysencks noch einmal übersichtlich dar:

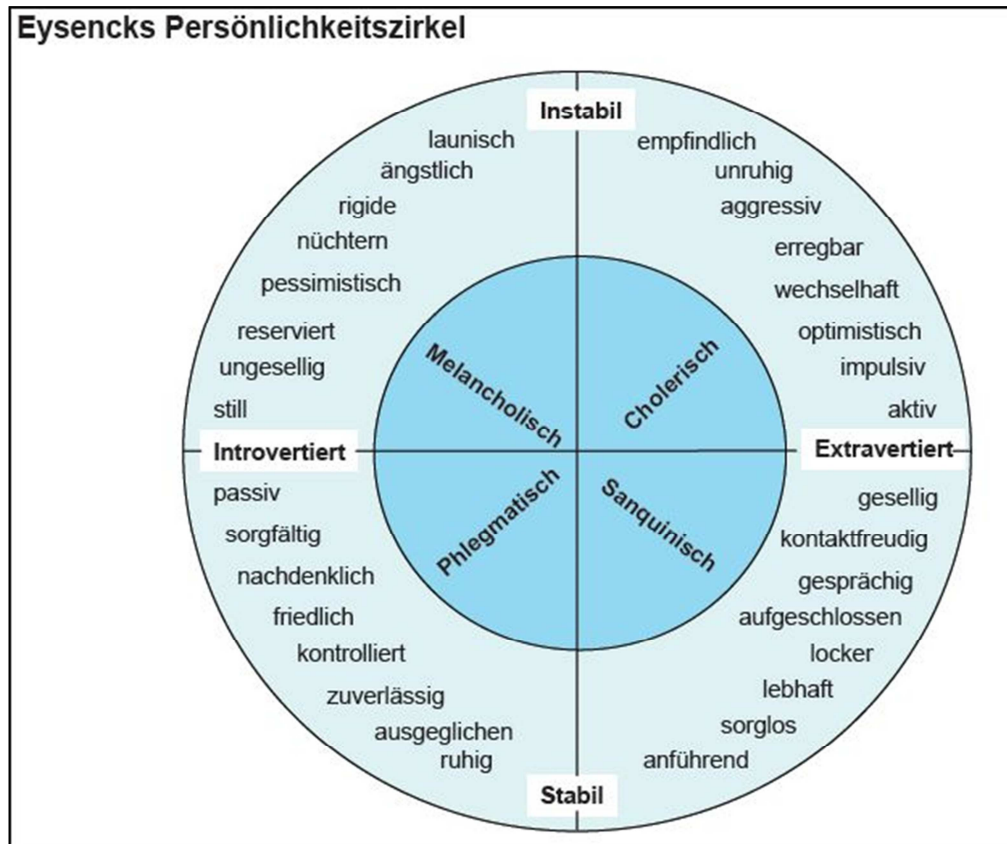


Abb. 9: Eysencks Persönlichkeitszirkel
(Quelle :<http://www.hipa.at/psycho/eigenschaftstheorien.htm>)

3.3.3 Die Entwicklung des Verfahrens

Der EPQ ist Teil einer Entwicklungslinie von Inventaren, die die Eysenckschen Persönlichkeitsdimensionen erfassen sollen.

Das erste in dieser Reihe war der Maudsley Medical Questionnaire (MMQ; Eysenck, 1952). Es umfasste 40 Fragen zur Erfassung von Neurotizismus bzw. Emotionalität und stellt den Ausgangspunkt seiner Fragebogenentwicklungen dar. Auf das MMQ folgte das Maudsley Personality Inventory (MPI; Eysenck, 1959, 1964). Eysenck fügte dem MPI eine zweite unabhängige Dimension, das Konstrukt der Extraversion hinzu. Damit entwickelte er das Eysenck Personality Inventory (EPI; Eysenck und Eysenck, 1964). Die deutsche Fassung des EPI wurde 1983 von Eggert veröffentlicht. Das EPI

enthielt eine zusätzliche L-Skala („Lie“), die zur Erfassung von Dissimulationen dienen sollte. Es kam in zwei Parallelförmigkeiten A und B vor. Durch das EPI wurden psychometrische Verbesserungen erzielt, wie z.B. die Steigerung der Reliabilität und die Beseitigung der negativen Korrelationen zwischen E und N.

Das Psychotizismus-Konstrukt, welches die bisher genannten Verfahren nicht beinhalten, wurde ursprünglich über objektive Tests gewonnen.¹⁵² Später wurden auch Fragebögen zu seiner Erfassung eingesetzt. Über das Studium zweier unveröffentlichter Inventare gelang man schließlich zum Eysenck Personality Questionnaire (EPQ; Eysenck & Eysenck, 1975), das aus 90 Fragen besteht. 25 dieser Items werden dabei der P-Skala, 21 der E-Skala, 23 der N-Skala und 21 der L-Skala zugerechnet. Um die Universalität der Konstrukte zu prüfen und eine Adaption des EPQ in unterschiedlichen Ländern zu ermöglichen wurden diese 90 Items und 11 nicht auswertbare Puffer-Items in viele Sprachen übersetzt (Barrett & Eysenck, 1984; Eysenck 1983).¹⁵³

Die P-Skala des EPQ und die damit verbundenen Mängel wurden häufig diskutiert (z.B.: Block, 1977). Speziell die Schiefe der Verteilung, die geringe Spannweite in den erzielten Ergebnissen und damit verbunden die geringe Differenzierungsfähigkeit und eine geringe interne Konsistenz wurden kritisiert. Eysenck und Eysenck (1991) stellten daraufhin eine 100 Fragen umfassende revidierte Form (EPQ-R) des EPQ vor. Die Items dieses Verfahrens verteilen sich wie folgt auf die vier Skalen: 32 P-Items, 23 E-Items, 24 N-Items und 21 L-Items, die sowohl negativ als auch positiv verschlüsselt sind.

Zuvor wurden zwei Stichproben, in denen neue P-Items getestet wurden, analysiert. Als Ergebnis wurden sechs der P-Items des EPQ aufgrund inkonsistenter Ladungen auf dem P-Faktor eliminiert. Die restlichen 19 Items wurden in die endgültige Form des EPQ-R übernommen. Dazu kamen 13 neue P-Items. Während auch den Skalen E und N je zwei Items bzw. ein Item hinzugefügt wurden, blieb die L-Skala unverändert.

Die Reliabilitäten (Cronbach's Alpha) sind mit .78 bei den Männern und .76 bei den Frauen, im Vergleich zum EPQ (.74 und .68) deutlich erhöht und hinreichend, erreichen aber trotzdem nicht die Werte der drei anderen Skalen (zwischen .79 und .90).¹⁵⁴

¹⁵² vgl. Ruch, Die revidierte Fassung des Eysenck Personality Questionnaire und die Konstruktion des deutschen EPQ-R bzw. EPQ-RK, 1999

¹⁵³ vgl. ebenda

¹⁵⁴ vgl. ebenda

Aus den Daten derselben Stichprobe ging auch eine Kurzversion des EPQ-R hervor (EPQ short scale). Diese besteht aus 48 Items, wobei diese sich gleichmäßig auf die vier Skalen verteilen.¹⁵⁵

Willibald Ruch wollte eine dem englischen EPQ-R möglichst vergleichbare deutsche Version schaffen. Auch eine deutsche Kurzform (EPQ-RK) sollte konstruiert werden. Um diesem Anspruch gerecht zu werden, übersetzte man die Items nicht nur einfach ins Deutsche. Vielmehr wurden die Items des englischen EPQ-R zum Beispiel ausführlichen Faktoranalysen unterzogen. Um die Skalen dann an das Deutsche zu adaptieren wurden Items ersetzt oder weggelassen. Vor allem der P-Skala wurde dabei besondere Beachtung geschenkt. Als Ergebnis der gesamten Prozedur liegt die Standardform des deutschen EPQ-R mit insgesamt 102 Fragen vor.¹⁵⁶

Mit Hilfe der Untersuchungsergebnisse des EPQ-R wurde dann die deutsche Kurzfassung EPQ-RK erstellt, welche insgesamt 50 Fragen enthält.

3.3.4 Skalen und Interpretationshinweise

Die Zusammenfassung der 50 Items des EPQ-RK ergibt die Rohwerte für vier Subskalen. Diese geben den individuellen Ausprägungsgrad der jeweils erfassten Faktoren der Persönlichkeit nach Eysenck wieder. In der folgenden Tabelle sind die vier Skalen mit ihren charakterisierenden Eigenschaften dargestellt:

Tab. 5: Die Skalen des EPQ-RK (nach Ruch; Online-Handbuch EPQ-R¹⁵⁷)

Skala	Charakterisierende Eigenschaften
Psychotizismus	Der Persönlichkeitstypus des Psychotizismus wird durch Eigenschaften wie aggressiv, kalt, egozentrisch, unpersönlich, impulsiv, antisozial, unempathisch, kreativ und hartherzig markiert.
Extraversion	Das Typenkonzept der Extraversion wird aus der Interkorrelation zwischen den Eigenschaften gesellig, lebhaft, aktiv, reizsuchend, sorglos, dominant, optimistisch und wagemutig abgeleitet.

¹⁵⁵ vgl. Ruch, Die revidierte Fassung des Eysenck Personality Questionnaire und die Konstruktion des deutschen EPQ-R bzw. EPQ-RK, 1999

¹⁵⁶ vgl. ebenda

¹⁵⁷ vgl. <http://www.unifr.ch/ztd/HTS/infest/WEB-Informationssystem/de/4dek01/d9ec3c7916bf11d58b620001028b2ad7/hb.htm>, verfügbar am 15.07.2010

Neurotizismus	Neurotizismus schließlich besteht aus Eigenschaften wie ängstlich, bedrückt, zu Schuldgefühlen neigend, geringe Selbstachtung, angespannt, irrational, schüchtern, launisch und emotional.
Lügenskala	Diese Skala dient der Kontrolle der Antworttendenzen in Richtung sozialer Erwünschtheit.

3.3.5 Normen und Gütekriterien

Zur Normierung und den Gütekriterien liegen aktuell noch keine ausreichenden Angaben vor. Eine Publikation ist in Vorbereitung.

Objektivität

Der standardisierte Fragebogen, der den Probanden jeweils zwei Antwortmöglichkeiten („ja“ und „nein“) vorgibt, gewährleistet eine relativ hohe Durchführungsobjektivität.

Die Auswertungsobjektivität wird durch die Anwendung eines vorgegebenen Auswertungsverfahrens gewährleistet. Dabei ist zu beachten, dass die Items unterschiedlich verschlüsselt sind. Eine entsprechende Tabelle liegt vor. Mit der Summierung der Werte in den einzelnen Subskalen erhält man die entsprechenden Rohwerte.

Die Interpretationsobjektivität des EPQ-RK leidet unter den dürftigen Angaben. Interpretationshinweise findet man weder in den diversen Veröffentlichungen, noch im Manual. Lediglich die Beschreibung der Skalen durch Adjektive kann zur Interpretation herangezogen werden.

Reliabilität

Die innere Konsistenz wird mit Cronbach's Alpha zwischen .72 und .85, die Retestreliabilität wird zwischen .76 und .90 angegeben.¹⁵⁸

¹⁵⁸ vgl. Ruch, Die revidierte Fassung des Eysenck Personality Questionnaire und die Konstruktion des deutschen EPQ-R bzw. EPQ-RK, 1999

Validität

Zur Validität lassen sich in der einschlägigen Literatur bisher noch keine Angaben finden. Das Testmanual weist auf eine Publikation hin, die sich noch in der Vorbereitung befindet.

Mit den Beschreibungen der Testverfahren ist der theoretische Teil dieser Arbeit abgeschlossen. Die folgenden Kapitel beschäftigen sich mit der empirischen Untersuchung der drei vorgestellten Persönlichkeitsinventare.

4 Empirische Belege und Fragestellungen

4.1 Empirische Belege

Die korrelativen Zusammenhänge, die im empirischen Teil der vorliegenden Arbeit untersucht werden sollen, waren bereits Gegenstand zahlreicher früherer Untersuchungen. Es soll überprüft werden, ob die Ergebnisse der früheren Studien repliziert werden können. Nachfolgend sind die für diese Arbeit relevanten Informationen aus den verschiedenen Studien zusammengefasst.

4.1.1 Die Neuinterpretation des Myers-Briggs Typenindikators aus der Perspektive des Fünf-Faktoren-Modells der Persönlichkeit

Diese ältere Studie aus dem Jahr 1989 wurde von Robert R. McCrae und Paul T. Costa, Jr. veröffentlicht. Sie beschäftigt sich mit der Neuinterpretation des Myers-Briggs Typenindikators (MBTI) aus der Perspektive des bekannten Fünf-Faktoren-Modells. An der Untersuchung nahmen 267 Männer und 201 Frauen teil. Sie vervollständigten jeweils den MBTI und das NEO-Persönlichkeitsinventar (NEO-PI).

Die zentrale Fragestellung dieser Untersuchung war, ob der MBTI tatsächlich die Typentheorie von Jung verkörpert, also ob dieses Verfahren seinem Anspruch gerecht wird, Typen zu messen und nicht Persönlichkeitseigenschaften oder ob es nicht eher aus der Sicht des Fünf-Faktoren-Modells, also aus traitorientierter Sicht, interpretiert werden sollte.

Das Fünf-Faktoren-Modell stellt eine alternative Basis dar, um die Ergebnisse des MBTI zu interpretieren. Die Autoren geben zu bedenken, dass das Konzept Jungs', teilweise sehr verdreht und verzerrt durch den MBTI repräsentiert wird. Viel von den Typenbeschreibungen von Jung bezieht sich auf das unbewusste Leben des Menschen. Dieses ist durch Selbsteinschätzung oft nicht erreichbar. Außerdem kritisieren sie, dass die Entwicklung der Skalen des MBTI ohne Beachtung der hypothetischen typologischen Struktur erfolgt ist.

Um eine Antwort auf die zentrale Fragestellung zu finden, wurden die erhobenen Daten analysiert. Man korrelierte u.a. die Faktoren des NEO-PI mit den einzelnen Skalen des MBTI und fand die gleichen Korrelationen wie in den zuvor beschriebenen

Studien. Damit gaben sich die Autoren nicht zufrieden und korrelierten die Faktoren des NEO-PI noch einmal mit den Gesamtskalen des MBTI – also EI, SN, TF und JP, um ein klareres Bild zu erhalten. Auch hier zeigten sich eindeutige Zusammenhänge zwischen dem Fünf-Faktoren-Modell und der Typologie.

Die Autoren kommen zusammenfassend zu dem Ergebnis, dass Jungs' Theorie zwar als Grundlage für den MBTI diene, aber nicht korrekt durch den MBTI operationalisiert wurde. Sie gehen davon aus, dass der MBTI kein Instrument ist, das die Typen von Jung misst, sondern ebenfalls Persönlichkeitseigenschaften erfasst – ähnlich wie das Fünf-Faktoren-Modell. Der MBTI weist beachtliche Zusammenhänge mit dem Fünf-Faktoren-Modell auf. Lediglich der Faktor Neurotizismus wird in diesem Verfahren nicht abgebildet.¹⁵⁹

Diese Studie, die das Konzept des MBTI in Frage stellt, ist Ausgangspunkt und Grundlage zahlreicher weiterer Untersuchungen bezüglich der Zusammenhänge zwischen dem Fünf-Faktoren-Modell und der Typologie. Das Fünf-Faktoren-Modell wird in den Studien durch unterschiedliche Instrumente repräsentiert.

4.1.2 Der Zusammenhang zwischen dem revidierten NEO-Persönlichkeitsinventar und dem Myers-Briggs Typenindikator

Die vorliegende Studie von Adrian Furnham, Joanna Moutafi und John Crump aus dem Jahr 2003 untersuchte die Korrelationen zwischen den Subfaktoren des Fünf-Faktoren-Modells (auch: Big-Five-Modell) und den Faktoren des Myers-Briggs-Typenindikator (MBTI). Als Vergleichsverfahren wurde in diesem Fall das revidierte NEO-Persönlichkeitsinventar (NEO-PIR) gewählt.

Die größte Überschneidung wurde zwischen der Dimensionen „Extraversion“ des NEO-PIR und der Skala „Extraversion-Introversion“ (kurz:EI) des MBTI festgestellt. Dieses Ergebnis wird mit der eindeutigen inhaltlichen Übereinstimmung der theoretischen Konstrukte begründet.

Weiterhin wurde festgestellt, dass der Faktor „Offenheit für neue Erfahrung“ des NEO-PIR negativ mit der Präferenz „Sensing“ (S) korreliert und als Folge dessen positiv mit der Präferenz „Intuition“ (N). S und N stellen die entgegengesetzten Präferenzen in

¹⁵⁹ vgl. McCrae/ Costa, 1989, S. 17 ff.

einer Dichotomie dar. Damit lässt sich diese „entgegengesetzte“ Korrelation erklären. N-Typen hinterfragen die Informationen, die durch ihre Sinne bereitgestellt werden, um nicht sofort augenscheinliche Möglichkeiten zu entdecken, während S-Typen die Informationen direkt durch ihre Sinne empfangen. Leider wird dieser Zusammenhang in der Studie nicht näher erklärt.

Der Faktor „Verträglichkeit“ des NEO-PIR korreliert negativ mit „Thinking“ (T) und positiv mit „Feeling“ (F). T und F sind ebenfalls entgegengesetzte Präferenzen in einer Dichotomie. T-Typen leben rational, analytisch, suchen nach logischen Zusammenhängen und analysieren Informationen über strikte Prinzipien und Kausalität. F-Typen dagegen sind mehr auf ihre Gefühle für eine Person oder ein Geschehen bezogen als durch logische Analyse herauszufinden, was sie von dieser Person oder diesem Geschehen halten. Somit wird auch hier die entgegengesetzte Korrelation durch die theoretische Konstruktion des MBTI und des NEO-PIR erklärt. Verträgliche Menschen neigen dazu, eher ein F-Typ zu sein.

Schließlich wurde noch eine Korrelation zwischen dem Faktor Gewissenhaftigkeit und den Präferenzen der vierten Dichotomie „Judging“ (J) und „Perceiving“ (P) gefunden. Gewissenhaftigkeit korreliert positiv mit J und negativ mit P. Gewissenhafte Menschen sind eher ordentlich und handeln überlegt. Die J-Typen werden im MBTI mit den gleichen Eigenschaften beschrieben. P-Typen sind eher ungestüm und impulsiv, was die negative Korrelation dieser beiden Faktoren erklärt.

Die Autoren weisen in dieser Studie lediglich die Zusammenhänge auf, lassen aber eine ausführliche Interpretation dieser vermissen.¹⁶⁰

4.1.3 Der Zusammenhang zwischen dem Myers-Briggs Typenindikator und dem Eysenck Personality Questionnaire unter erwachsenen Kirchgängern

Leslie J. Francis und Susan H. Jones untersuchten 2000 den Zusammenhang zwischen dem Myers-Briggs Typenindikator (MBTI) und dem Eysenck Personality Questionnaire (EPQ). Der Kontext der Untersuchung bleibt für diese Arbeit unberücksichtigt. 377 Teilnehmer vervollständigten im Zuge dieser empirischen Untersuchung den MBTI und den EPQ.

¹⁶⁰ vgl. Furnham/ Moutafi/ Crump, 2003

Francis und Jones stellen in dieser Studie fest, dass sich die MBTI-Skalen durchaus vollständig auf das dreidimensionale Modell des EPQ abbilden lassen. Die Messungen des Faktors Extraversion aus dem EPQ korrelieren hoch mit der Extraversion-Introversion-Dichotomie (EI) des MBTI. Dies impliziert, dass beide unabhängige Methoden die gleichartigen Konstrukte messen. Weiterhin korreliert die Präferenz Extraversion aus dem MBTI signifikant aber nicht stark negativ mit dem Faktor Neurotizismus und positiv mit dem Faktor Psychotizismus des EPQ. Folgerichtig konstruiert der MBTI die Extravertierten als stabiler und die Introvertierten als weniger stabil. Die positive Korrelation mit Psychotizismus deutet an, dass der MBTI die Extravertierten als hart und kompromisslos und die Introvertierten als zartbesaitet konzipiert. Es lassen sich mäßige Korrelationen zwischen dem Psychotizismus-Konstrukts des EPQ und der „Perceiving-Judging“-Dichotomie feststellen. Man kann schlussfolgern, dass die P-Typen eher als kompromisslos und die J-Typen als zartbesaitet im MBTI ersinnt werden. Das MBTI-Konzept charakterisiert die „Sensing“-Typen im Gegensatz zu den intuitiven Typen als mehr introvertiert und weniger stabil (MBTI-SN korreliert signifikant, aber nicht hoch mit EPQ-Extraversion und EPQ-Psychotizismus). Weiterhin treten signifikante Korrelationen zwischen der „Thinking-Feeling“-Dichotomie und dem Faktor Neurotizismus des EPQ auf. Es wird angedeutet, dass F-Typen eher weniger stabil sind als T-Typen.

Die Autoren ziehen wichtige Schlussfolgerungen aus den skizzierten Ergebnissen.

Die Untersuchungen untermauern die Ergebnisse der früheren Studien (z.B. Steele und Kelly, 1976; Landrum, 1992). Sowohl die MBTI-Messungen als auch die EPQ-Messungen weisen beide ein hohes Maß an Varianz auf. Obwohl die zugrundeliegenden Modelle verschiedene Definitionen von Intro- und Extraversion vorgeben, scheinen beide einen ähnlichen Aspekt der Persönlichkeit zu berühren.

Die Daten zeigen eine ziemlich hohe Korrelation zwischen der EPQ-P-Skala und der MBTI-JP-Dichotomie. Diese Beziehung ist insofern widerspruchsfrei, dass beide – also die EPQ-P-Skala und die P-Präferenz des MBTI eine wesentliche Komponente von Impulsivität enthalten. Daher erscheint es auch vernünftig die Forschungsergebnisse beider Methoden zu vergleichen.

Im Gegensatz zu früheren Ergebnissen von Wakefield, Sasek, Brabeker und Friedmann (1976) weisen die vorliegenden Daten darauf hin, dass die MBTI-Skalen tatsächlich in einer komplexen Art und Weise auf das drei-dimensionale Modell des EPQs abgebildet werden können.

4.1.4 Die Untersuchung zwischen drei Verfahren der Selbsteinschätzung bezüglich Persönlichkeit, Typen und Störungen

Diese Studie wurde 2005 von Adrian Furnham und John Crump vorgelegt. An der Untersuchung haben 858 Probanden teilgenommen. Sie vervollständigten alle den Hogan Development Survey (HDS) von Hogan und Hogan (1997), der elf Verhaltenstendenzen, die Karrieren zum Scheitern bringen sollen, misst.¹⁶¹ Da dieser für die vorliegende Arbeit nicht relevant ist, wird nicht näher auf den Inhalt des Verfahrens und die dazu vorliegenden Ergebnisse eingegangen. Weiterhin bearbeiten alle Teilnehmer das NEO- Persönlichkeitsinventar (NEO-PI). Nur 431 Probanden nahmen an allen drei Verfahren teil.

Zunächst greifen Furnham und Crump die Kritikpunkte von McCrae und Costa (siehe Abschnitt 4.1.1) am MBTI auf. Sie kritisieren ebenfalls die fehlende Struktur im Konzept von Jung und benennen einige Studien in denen die Validität des Konzeptes nicht bestätigt werden konnte.

Als für diese Arbeit relevantes Ergebnis lässt sich nur eine Aussage festhalten. Auch in dieser Studie konnte der starke Zusammenhang zwischen dem Faktor Extraversion im NEO-PI und der Präferenz Extraversion des MBTI nachgewiesen werden. Auf weitere Korrelationen zwischen den beiden Verfahren wird weder in der Ergebnisdarstellung noch in der Diskussion eingegangen.¹⁶²

4.2 Zentrale Fragestellungen

Die empirische Untersuchung in der vorliegenden Arbeit beschäftigt sich mit dem korrelativen Vergleich von drei ausgewählten Persönlichkeitsverfahren. Diese drei Verfahren basieren auf unterschiedlichen Modellen der Persönlichkeit. Aus den dargestellten Studien, die sich bereits mit Verfahren beschäftigt haben, denen dieselben Persönlichkeitsmodelle zugrunde liegen, ergeben sich für diese Arbeit zwei zentrale Fragestellungen.

Zunächst sollen die Verfahren einzeln untersucht werden. Es geht darum herauszufinden, ob Korrelationen innerhalb der Verfahren auftauchen. Ist dies der Fall,

¹⁶¹ vgl. http://www.metaberatung.de/05_Tests_und_Methoden/HDS.htm, verfügbar am 03.08.2010

¹⁶² vgl. Furnham/ Crump, 2005, S. 167 ff.

bedeutet das, dass ein Merkmal durch ein anderes mitgemessen wird. Die Konstrukte sind in so einem Fall nicht „sauber“ getrennt und es existieren Überschneidungen. Für die Güte eines Testverfahrens hat dies keinen erheblichen Einfluss. Bestehende Korrelationen innerhalb eines Verfahrens können jedoch unter Umständen konvergente oder diskriminante Validitäten mit einem anderen Verfahren erklären.

Die konvergenten und diskriminanten Validitäten der drei Verfahren sollen in einem zweiten Schritt untersucht werden. Dazu werden die Verfahren nacheinander miteinander korreliert. Ziel ist es dabei, Übereinstimmungen zwischen den Verfahren und damit den zugrundeliegenden Persönlichkeitsmodellen zu finden und durch die theoretische Konzeption des jeweiligen Verfahrens erklären zu können.

Es soll überprüft werden, ob die Ergebnisse aus den vorgestellten Studien, die ausschließlich im angelsächsischen Sprachraum durchgeführt worden, mit der vorliegenden deutschen Stichprobe der Hochschule Mittweida, repliziert werden können.

5 Methodisches Vorgehen

Nachfolgend soll ein Überblick über den Inhalt und Ablauf der Untersuchung gegeben werden. Dieser beinhaltet die Beschreibung der Stichprobe, den Ablauf der Datenerhebung, die entsprechende Aufbereitung der Daten, sowie die gewählten Analyseverfahren, um die Daten entsprechend auszuwerten.

5.1 Stichprobenbeschreibung

Die Stichprobe besteht aus insgesamt 100 Studenten ($N=100$), die im Rahmen einer Veranstaltungsreihe an der Fachhochschule Mittweida ausgewählte Testverfahren absolvierten. An der Untersuchung nahmen 55 männliche und 45 weibliche Probanden teil. Das Durchschnittsalter beträgt 25,1 Jahre, wobei der jüngste Untersuchungsteilnehmer 18,9 Jahre und der älteste 40,2 Jahre alt sind. Zum Zeitpunkt der Untersuchungen besitzen die meisten Teilnehmer ($N=91$) das Abitur. Jeweils vier Probanden haben die Fachhochschule bzw. die Universität bereits abgeschlossen. Ein Teilnehmer besitzt den Realschulabschluss. Es konnten alle Datensätze für die vorliegende empirische Untersuchung genutzt werden.

5.2 Datenerhebung

Die drei untersuchten Verfahren fanden immer in einer feststehenden, aber unwillkürlich gewählten, Reihenfolge statt. Die Teilnehmer wurden gebeten, zunächst die computergestützte Version des NEO-FFI zu bearbeiten. Danach folgten der EPQ-RK und der MBTI in Papierform. Um konzentriertes Arbeiten zu ermöglichen und äußere Einflüsse zu minimieren, wurden die Probanden einzeln getestet und jeweils von einem geschulten Testleiter begleitet. Zudem fanden die Untersuchungen in einem Büro der Fachhochschule Mittweida statt, um weitgehend gleiche Untersuchungsbedingungen zu gewährleisten. Die Datenerhebung erfolgte im Zeitraum vom November 2007 bis November 2009. Alle Untersuchungsteilnehmer wurden über den Ablauf der Untersuchung und die Speicherung der Daten informiert. Weiterhin erteilten alle Probanden schriftlich ihr Einverständnis, dass ihre Daten gespeichert und für wissenschaftliche Arbeiten und Untersuchungen genutzt werden dürfen.

Vor der Bearbeitung des computergestützten NEO-Fünf-Faktoren-Inventars wurde den Teilnehmern der Umgang mit der Software („Hogrefe-Testsystem“) und dem Eingabegerät (Maus) erläutert und damit verbundene Fragen beantwortet. Weiterhin wies man sie noch einmal darauf hin, dass es weder eine Zeitbeschränkung bei der Bearbeitung der Aufgaben, noch falsche oder richtige Antworten gibt. Auf weitere Bearbeitungshinweise wurde aus Gründen einer objektiven Durchführung weitestgehend verzichtet. Die Bearbeitungszeit dieses Verfahrens lag durchschnittlich bei 8,7 Minuten.

Der darauf folgende EPQ-RK wurde als Paper-Pencil-Test durchgeführt. Die Probanden bekamen den Fragebogen mit der Bitte vorgelegt, die Anleitung gründlich zu lesen und nachfolgend die Fragen zu beantworten. Nach der Beantwortung der Fragen durch die Teilnehmer wurde der Fragebogen durch den Testleiter kontrolliert. Aufgrund des dichotomen Antwortformates des EPQ-RK neigten einige Teilnehmer dazu, beide Antworten anzukreuzen oder das Kreuz genau zwischen die Antwortmöglichkeiten zu setzen. In diesem Fall wurde der Fragebogen dem Probanden noch einmal zur Korrektur der betreffenden Items vorgelegt.

Der MBTI bildete den Abschluss der Testreihe und wurde den Teilnehmern ebenfalls als Paper-Pencil-Test vorgelegt. Nachdem sie die Anleitung durchgelesen hatten und eventuelle Verständnisfragen geklärt waren, begannen die Probanden mit der Bearbeitung der Fragen. Auch dieses Verfahren arbeitet, ähnlich wie der EPQ-RK, mit einem dichotomen Antwortformat. Damit traten die bereits geschilderten Probleme auch hier auf. Die Probanden wurden ebenfalls um Nachkorrektur gebeten.

5.3 Datenaufbereitung

Für die Aufbereitung und Auswertung der erhobenen Daten wurde ein Datenanalyseprogramm von IBM® SPSS® Statistics (kurz: SPSS) genutzt. Predictive Analytics Software Statistics 18 (kurz: PASW) ermöglicht dem Nutzer eine strukturierte Erfassung seiner statistischen Daten, sowie eine korrekte Analyse und ordentliche Aufbereitung seiner Untersuchungsergebnisse.

Die Daten des computergestützten Verfahrens (NEO-FFI) wurden nach jeder Testung automatisch gespeichert. Die Software („Hogrefe-Testsystem“) ermöglicht es, die im Testsystem gespeicherten Daten in andere Programme zu exportieren. Dies hat den

Vorteil, dass die Daten ohne Verluste und Formatierungsprobleme in einem Datenanalyseprogramm wie PASW geöffnet und bearbeitet werden können. Die Daten wurden also in das Statistikprogramm exportiert, die Datensätze geordnet und nicht benötigte Daten wie z.B. Normwerte gelöscht, um ein übersichtliches Datentableau zu erhalten. Dieses beinhaltete alle relevanten Daten des NEO-FFI.

Die Daten des EPQ-RK und des MBTI mussten manuell in das Datenanalyseprogramm PASW eingetragen werden, da beide Verfahren nur in der Paper-Pencil-Test-Version vorlagen. Dies ist insoweit nachteilig, dass während der manuellen Eingabe der Daten Fehler gemacht werden können, die sich möglicherweise auf die Ergebnisse auswirken. Die Vor- und Nachteile von Paper-Pencil sowie computergestützten Testverfahren werden u.a. bei Fisseni (2004) oder Hänsgen (1999) ausführlich erläutert.

Nachdem die Daten der beiden Testverfahren „per Hand“ in das PASW-System eingepflegt wurden, konnten sie mit der bereits bestehenden Datei des NEO-FFI verknüpft werden, so dass eine Datei mit allen relevanten Daten entstand. Diese wurde nochmals strukturiert und konnte dann für die Datenauswertung verwendet werden.

5.4 Datenauswertung

5.4.1 Test auf Normalverteilung

Auf der Grundlage der aufbereiteten Daten können Berechnungen und Vergleiche hinsichtlich der zentralen Fragestellungen aus dem Abschnitt 4.2 durchgeführt werden.

Verschiedene Verfahren der Statistik sind nur dann sinnvoll anwendbar, wenn annähernd eine Normalverteilung der Daten vorliegt. Korrelationen oder auch arithmetische Mittel sind wenig aussagekräftig, wenn die Verteilung der vorliegenden Daten durch Ausreißer oder extreme Schiefen geprägt sind. Aus diesem Grund wird auch in der vorliegenden Arbeit ein Test auf Normalverteilung durchgeführt. Für den Nachweis einer Normalverteilung kann grundsätzlich auf verschiedene Methoden zurückgegriffen werden. Für eine optische Abschätzung der Normalverteilung kann die grafische Darstellung genutzt werden. Mit den Kolmogorov-Smirnov-Test und dem Chi-Quadrat-Test stehen jedoch auch zwei mathematisch-statistische Methoden zur Verfügung, um eine Normalverteilung festzustellen. Da eine grafische Betrachtung mitunter keine zuverlässigen Ergebnisse liefert, greift die Autorin auf den Kolmogorov-

Smirnov-Test zurück. Dieser ist ein nicht-parametrischer Test, der auch bei kleinen Stichproben eingesetzt werden kann, um zu prüfen, ob eine gegebene Verteilung mit hoher Wahrscheinlichkeit von der Normalverteilung abweicht. Die Berechnung basiert auf dem Vergleich der Verteilung der vorliegenden Daten mit einer hypothetischen Normalverteilung. Dabei wird davon ausgegangen, dass die tatsächliche Verteilung von einer hypothetischen Normalverteilung an einem beliebigen Punkt eine bestimmte flächenmäßige Abweichung nicht überschreiten darf, andernfalls muss die Annahme einer Normalverteilung verworfen werden. In einer Reihe von Rechenschritten kann die jeweilige konkrete Abweichung errechnet werden. Mit PASW können die Werte natürlich wesentlich komfortabler ermittelt werden.¹⁶³

5.4.2 Korrelationsanalysen

Um die zentralen Fragestellungen dieser Arbeit beantworten zu können, ist es notwendig festzustellen, ob überhaupt Zusammenhänge zwischen den drei ausgewählten Verfahren bestehen. Um derartige Zusammenhänge aufzudecken, gibt es verschiedene Methoden. Die Analyse von Korrelationen ist dabei wohl eines der gebräuchlichsten Verfahren und soll auch im empirischen Teil der vorliegenden Arbeit genutzt werden.

Die Korrelationsanalyse ist ein hypothesenprüfendes Verfahren, dass benutzt wird, um Zusammenhänge zwischen zwei oder mehreren Variablen zu berechnen. Grundsätzlich können folgende Zusammenhänge bestehen:

☛ Übereinstimmung:

Je höher der Wert der Variablen A ist, desto höher ist meist auch der Wert der Variablen B. Man spricht in diesem Fall von einer *positiven Korrelation*.

☛ Gegensatz:

Je höher der Wert der Variable A ist, desto niedriger ist meist der Wert der Variable B. Man bezeichnet das häufig auch als *negative Korrelation*.

¹⁶³ vgl. <http://www.univie.ac.at/ksa/elearning/cp/quantitative/quantitative-59.html>, verfügbar am 03.08.2010

■ Unabhängigkeit:

In diesem Fall besteht *keine Korrelation*. Hohe Werte der Variable A können relativ beliebigen Werten der Variablen B entsprechen und umgekehrt.¹⁶⁴

Der Zusammenhang wird vom sogenannten Korrelationskoeffizienten beschrieben, der Werte zwischen -1,00 und +1,00 annehmen kann. Das Vorzeichen gibt an, ob der Zusammenhang gleichläufig (positiv) oder gegenläufig (negativ) ist. Ein Wert von -1,0 bedeutet daher eine perfekte negative Korrelation: Hohe Werte der Variablen A gehen also ausnahmslos mit niedrigen Werten der Variablen B einher und umgekehrt. Ein Wert von (+)1,00 dagegen bezeichnet eine perfekte positive Korrelation, d.h. hohe Werte der Variablen A entsprechen immer hohen Werten der Variablen B und umgekehrt.

Gefundene Korrelationen sollten nochmals kritisch hinterfragt werden. Es gibt u.a. die Scheinkorrelation. Diese treten nur deshalb auf, weil beide Variablen hoch mit einer dritten Variablen korrelieren. Weiterhin gibt es das Phänomen der verdeckten Korrelation. Davon spricht man, wenn statistisch keinerlei Korrelationen errechnet wurden, aber sachlich eindeutig ein Zusammenhang besteht. Hier neutralisieren sich die Subgruppen der vorliegenden Stichprobe gegenseitig, weil sie sich gegenläufig verhalten.¹⁶⁵

Die Aussagekraft einer Korrelation hängt von

- der Höhe der Korrelation
- der Größe der Stichprobe und
- der Sorgfalt beim Ausschluss einer möglichen Scheinkorrelation und des Erkennens verdeckter Korrelationen

ab.

Die statistische Kennzahl der Signifikanz bezeichnet die Wahrscheinlichkeit eines statistischen Zusammenhanges zwischen den Variablen und berücksichtigt diese genannten drei Faktoren. Sie drückt aus, ob ein scheinbarer Zusammenhang rein zufälliger Natur sein könnte oder mit hoher Wahrscheinlichkeit tatsächlich vorliegt. Man spricht bei der Signifikanz von Irrtumswahrscheinlichkeiten oder Signifikanzniveaus. Ist eine Stichprobe sehr klein, muss die Korrelation extrem groß ausfallen, um überhaupt

¹⁶⁴ vgl. <http://www.univie.ac.at/ksa/elearning/cp/quantitative/quantitative-96.html>, verfügbar am 03.08.2010

¹⁶⁵ vgl. <http://www.univie.ac.at/ksa/elearning/cp/quantitative/quantitative-96.html>, verfügbar am 03.08.2010

signifikant sein zu können. Dagegen kann auch eine geringere Korrelation von 0.2 bei sehr großen Stichproben signifikant werden.¹⁶⁶

Die Signifikanz von Korrelationen wird von Statistikprogrammen wie PASW automatisch ermittelt. Es wird mit Sternen ausgedrückt, ob eine Korrelation signifikant ist, d.h. ob der Zusammenhang weitgehend gesichert erscheint.¹⁶⁷

Inzwischen haben drei wesentliche Schwellen der Irrtumswahrscheinlichkeit p breite Akzeptanz gefunden:

- $p \leq 0,05$: Die Wahrscheinlichkeit einer rein zufälligen Korrelation liegt bei unter 5 %, d.h. die Irrtumswahrscheinlichkeit ist kleiner als 5 %. In diesem Fall vergibt PASW einen Stern (*).
- $p \leq 0,01$: Die Irrtumswahrscheinlichkeit ist kleiner als 1 %. PASW vergibt in diesem Fall zwei Sterne (**).
- $p \leq 0,001$: Die Irrtumswahrscheinlichkeit ist kleiner als 1 ‰. Dann vergibt PASW drei Sterne (***).

¹⁶⁶ vgl. <http://www.univie.ac.at/ksa/elearning/cp/quantitative/quantitative-108.html>,
verfügbar am 04.08.2010

¹⁶⁷ Es werden 1, 2 oder 3 Sterne vergeben.

6 Ergebnisdarstellung und Interpretation

6.1 Test auf Normalverteilung

Nachfolgend sind die Ergebnisse des Kolmogorov-Smirnov-Tests für alle drei Verfahren, die mithilfe von PASW durchgeführt wurden, tabellarisch dargestellt:

Tab. 6: Prüfung der Daten des NEO-FFI auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest_NEO-FFI						
		Neurotizismus	Extraversion	Offenheit für Erfahrung	Verträglichkeit	Gewissenhaftigkeit
N		100	100	100	100	100
Parameter der Normalverteilung ^{a, b}	Mittelwert	1,44840	2,5288	2,32467	2,6017	2,8269
	Standardabweichung	,551285	,48716	,480254	,45318	,52124
Extremste Differenzen	Absolut	,095	,142	,076	,068	,062
	Positiv	,095	,069	,076	,049	,060
	Negativ	-,047	-,142	-,063	-,068	-,062
Kolmogorov-Smirnov-Z		,950	1,419	,763	,683	,617
Asymptotische Signifikanz (2-seitig)		,327	,036	,605	,740	,841
a. Die zu testende Verteilung ist eine Normalverteilung.						
b. Aus den Daten berechnet.						

Tab. 7: Prüfung der Daten des EPQ-RK auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest_EPQ-RK				
		Neurotizismus	Psychotizismus	Extraversion
N		100	100	100
Parameter der Normalverteilung ^{a, b}	Mittelwert	3,6200	2,9000	8,1900
	Standardabweichung	2,54963	1,71447	3,46904
Extremste Differenzen	Absolut	,151	,180	,159
	Positiv	,151	,180	,136
	Negativ	-,082	-,094	-,159
Kolmogorov-Smirnov-Z		1,508	1,802	1,591
Asymptotische Signifikanz (2-seitig)		,021	,003	,013
a. Die zu testende Verteilung ist eine Normalverteilung.				
b. Aus den Daten berechnet.				

Tab. 8: Prüfung der Daten des MBTI auf Normalverteilung

Kolmogorov-Smirnov-Anpassungstest_MBTI									
		MBTI_E	MBTI_I	MBTI_S	MBTI_N	MBTI_T	MBTI_F	MBTI_J	MBTI_P
N		69	31	61	39	74	26	65	35
Parameter der Normalverteilung ^{a,b}	Mittelwert	29,8116	26,8710	16,4918	14,7949	24,8784	13,4615	26,9077	20,9429
	Standardabweichung	16,85690	16,55645	11,16038	10,78786	16,62145	7,74199	14,26723	14,32797
Extremste Differenzen	Absolut	,085	,109	,121	,176	,115	,144	,069	,147
	Positiv	,073	,109	,121	,176	,115	,144	,069	,147
	Negativ	-,085	-,091	-,083	-,137	-,075	-,117	-,054	-,096
Kolmogorov-Smirnov-Z		,703	,604	,946	1,099	,987	,735	,560	,867
Asymptotische Signifikanz (2-seitig)		,706	,858	,332	,179	,285	,652	,912	,440
a. Die zu testende Verteilung ist eine Normalverteilung.									
b. Aus den Daten berechnet.									

Anmerkung:

E – Extraversion, I – Introversion, S – Sensing, N – Intuition, T – Thinking,

F – Feeling, J – Judging, P – Perceiving

Aus den von PASW ausgegebenen Tabellen ist noch nicht ersichtlich, ob eine Normalverteilung angenommen werden kann. Der Parameter a (in den Tabellen blau markiert) präsentiert nicht das Ergebnis, sondern die Nullhypothese des Tests, d.h. es wird eine Normalverteilung angenommen. Diese kann jedoch immer noch verworfen werden. Für die Überprüfung der Nullhypothese sind die rot markierten Werte in den Tabellen interessant. Diese werden mit einer entsprechenden Tabelle für den Kolmogorov-Smirnov-Test verglichen.¹⁶⁸ Auf eine ausführliche Darstellung des Ablaufes wird aus Gründen der besseren Lesbarkeit und der Komplexität des Verfahrens an dieser Stelle verzichtet.

Für alle drei Verfahren konnte die Nullhypothese nach dem Vergleich der relevanten Werte angenommen werden, d.h. es kann davon ausgegangen werden, dass eine Normalverteilung vorliegt.

Es ist jedoch zu beachten, dass der Kolmogorov-Smirnov-Test (v.a. bei kleinen Stichproben), extreme Abweichungen von einer Normalverteilung benötigt, um auf höheren Signifikanzniveaus die Annahme der Normalverteilung zu verwerfen. Ergo ist eine Nichtverwerfung der Annahme einer Normalverteilung noch kein Beweis für das Vorliegen einer solchen. Theoretisch kann man jedoch von einer Normalverteilung

¹⁶⁸ Entsprechende Tabellen sind sowohl in Formelsammlungen als auch im Internet zu finden. (u.a. http://www.statistik.tuwien.ac.at/public/dutt/vorles/mb_wi_vt/node98.html, verfügbar am 23.07.2010)

ausgehen. Nachfolgend wird daher eine Normalverteilung der vorliegenden Daten angenommen. Auf dieser Basis werden die folgenden Korrelationsanalysen durchgeführt.

6.2 Interkorrelationen in den einzelnen Verfahren

Zunächst werden, wie in Abschnitt 4.2 beschrieben, die Korrelationen der Verfahren getrennt betrachtet.

Aus Gründen der besseren Überschaubarkeit wurden die Tabellen bearbeitet und nur die relevanten Informationen dargestellt. Die signifikanten Korrelationen wurden blau hinterlegt. Die Originaltabellen aus dem Statistikprogramm PASW sind im Anhang zu finden.

NEO-Fünf Faktoren-Inventar (NEO-FFI)

Bei der Betrachtung der Korrelationen des NEO-FFI fallen insgesamt vier relevante Zusammenhänge auf.

Das Merkmal Extraversion korreliert mit $-.478$ hoch signifikant mit dem Merkmal Neurotizismus. Das negative Vorzeichen bedeutet, dass der Zusammenhang negativer Natur ist. Das bedeutet, dass extravertierte Menschen eher weniger neurotisch, d.h. emotional instabil, sind. Diese beiden Merkmale scheinen sich teilweise zu überschneiden. Neurotische Personen werden im NEO-FFI u.a. als unsicher und verlegen beschrieben. Sie machen sich eher Sorgen um ihre Zukunft, während Extravertierte als gesellig, aktiv und optimistisch charakterisiert werden. Die gegenläufige, also negative Korrelation erklärt sich damit aus der Konstruktion der beiden Merkmalsbereiche.

Der Faktor Verträglichkeit korreliert signifikant aber nur gering mit den Faktoren Neurotizismus und Extraversion. Die negative Korrelation zwischen Neurotizismus und Verträglichkeit bedeutet, dass neurotische Menschen eher weniger verträglich sind. Verträgliche Menschen werden im Manual als u.a. als Menschen beschrieben, die zu zwischenmenschlichem Vertrauen neigen. Eine neurotische Person wird eben diese Eigenschaft wahrscheinlich eher nicht haben. Sie neigen eher zu Unsicherheit und

Ängstlichkeit. Die positive Korrelation von Verträglichkeit mit Extraversion ist ebenfalls ein zu erwartender Zusammenhang. Beide Merkmale beschreiben Menschen als herzlich und kooperativ.

In der folgenden Tabelle sind die Interkorrelationen im NEO-FFI abzulesen. Es fällt eine vierte Korrelation zwischen dem Merkmal Gewissenhaftigkeit und Verträglichkeit auf, die nicht unbedingt zu erwarten war.

Tab. 9: Die Interkorrelationen im NEO-FFI

		Neuroti- zismus	Extra- version	Offenheit für Erfahrung	Verträglich- keit	Gewissen- haftigkeit
Neuroti- zismus	Korrelation nach Pearson	1				
	Signifikanz (2-seitig)					
	N	100				
Extra- version	Korrelation nach Pearson	-,478**	1			
	Signifikanz (2-seitig)	,000				
	N	100	100			
Offenheit für Erfahrung	Korrelation nach Pearson	-,053	,114	1		
	Signifikanz (2-seitig)	,602	,260			
	N	100	100	100		
Verträglich- keit	Korrelation nach Pearson	-,200*	,233*	-,027	1	
	Signifikanz (2-seitig)	,046	,020	,793		
	N	100	100	100	100	
Gewissen- haftigkeit	Korrelation nach Pearson	-,129	,001	-,135	,275**	1
	Signifikanz (2-seitig)	,202	,990	,181	,006	
	N	100	100	100	100	100

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

* . Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Um die geringe aber signifikante Korrelation zwischen Gewissenhaftigkeit und Verträglichkeit zu erklären, sollen auch hier die Definitionen beider Merkmale genauer

betrachtet werden. Die Korrelation würde bedeuten, dass gewissenhafte Personen auch verträglicher sind.

Verträgliche Personen werden als altruistisch, mitfühlend, verständnisvoll beschrieben. Sie neigen zu zwischenmenschlichem Vertrauen, Kooperativität und zur Nachgiebigkeit. Außerdem besitzen sie ein starkes Harmoniebedürfnis. Die Skala Gewissenhaftigkeit unterscheidet ordentliche, zuverlässige, hart arbeitende, disziplinierte, pünktliche, penible und ehrgeizige von nachlässigen und gleichgültigen Personen.¹⁶⁹ Aus diesen Beschreibungen lässt sich kein offener Zusammenhang ableiten. Eine direkte Überschneidung der beiden Merkmale durch gemeinsame Eigenschaften gibt es nicht.

In anderen vergleichbaren Studien oder den bekannten Normstichproben ist diese Korrelation nicht nachgewiesen worden.¹⁷⁰ Somit kann es sich um ein mögliches Spezifikum dieser Stichprobe handeln.

Eysenck Personality Questionnaire (EPQ-RK)

In der folgenden Tabelle sind die Interkorrelationen im EPQ-RK dargestellt.

Tab. 10: Die Interkorrelationen im EPQ-RK

		Neurotizismus	Psychotizismus	Extraversion
Neurotizismus	Korrelation nach Pearson	1		
	Signifikanz (2-seitig)			
	N	100		
Psychotizismus	Korrelation nach Pearson	,054	1	
	Signifikanz (2-seitig)	,596		
	N	100	100	
Extraversion	Korrelation nach Pearson	-,188	,129	1
	Signifikanz (2-seitig)	,061	,201	
	N	100	100	100

¹⁶⁹ vgl. Borkenau/ Ostendorf, 1993, S. 5

¹⁷⁰ vgl. Schumann, 2005, S. 37

Aus der Tabelle ist ersichtlich, dass in diesem Verfahren keine Interkorrelationen zwischen den Merkmalen bestehen. Eysenck verfeinert in der revidierten Form des EPQ das Merkmal Extraversion über die Komponente der Impulsivität und stimmte das Merkmal ganzheitlich mit dem neueren Faktor Psychotizismus ab. Dass keine Interkorrelationen in dem Verfahren auftreten könnte eine Folge dieser Verfeinerung sein, denn für andere Fragebögen, denen ebenfalls das Persönlichkeitsmodell Eysenck's als Grundlage dient, wurden solche Interkorrelationen berichtet.

Da innerhalb des Verfahrens keine Korrelationen auftreten, kann man davon ausgehen, dass die Merkmale sich nicht überschneiden.

Myers-Briggs-Typenindikator (MBTI)

Bei der Betrachtung der Tabellen in den nächsten Abschnitten fällt auf, dass der Stichprobenumfang N bezüglich der gesamten Skalen des MBTI nie 100 beträgt. Dies ist der Methodik der Datenaufbereitung und Datenauswertung geschuldet und keine fehlerhafte Darstellung. In der folgenden Tabelle sind die einzelnen Präferenzen des MBTI aufgelistet. Darunter stehen jeweils gültige und fehlende Werte. In diesem Verfahren existieren vier dichotome Skalen mit jeweils zwei Präferenzen. Der Proband wird auf einer Skala nur einer Präferenz zugeordnet. Damit wird die andere ausgeschlossen. Wenn also eine Person extravertiert ist, kann sie nicht gleichzeitig introvertiert sein. Da das Verfahren nicht dimensionsweise betrachtet wird, sondern die einzelnen Präferenzen zur Auswertung herangezogen werden, entstehen die fehlenden Werte in der abgebildeten Tabelle. Die Spaltensummen ergeben jeweils die Größe der Gesamtstichprobe $N = 100$.

Tab. 11: Verteilung der Präferenzen des MBTI

		E	I	S	N	T	F	J	P
N	Gültig	69	31	61	39	74	26	65	35
	Fehlend	31	69	39	61	26	74	35	65

Folgerichtig kann man aus den gültigen Werten auch die Verteilung der einzelnen Präferenzen über die gesamte Stichprobe entnehmen. Anschließend sind die Interkorrelationen im MBTI ebenfalls tabellarisch dargestellt.

Tab. 12: Die Interkorrelationen im MBTI

		E	I	S	N	T	F	J	P
E	Korrelation nach Pearson	1							
	Signifikanz (2-seitig)								
	N	69							
I	Korrelation nach Pearson	. ^a	1						
	Signifikanz (2-seitig)	.							
	N	0	31						
S	Korrelation nach Pearson	,002	-,150	1					
	Signifikanz (2-seitig)	,989	,540						
	N	42	19	61					
N	Korrelation nach Pearson	,077	-,358	. ^a	1				
	Signifikanz (2-seitig)	,704	,254	.					
	N	27	12	0	39				
T	Korrelation nach Pearson	,298 [*]	-,084	,084	-,241	1			
	Signifikanz (2-seitig)	,034	,703	,562	,257				
	N	51	23	50	24	74			
F	Korrelation nach Pearson	-,118	,196	,334	-,015	. ^a	1		
	Signifikanz (2-seitig)	,641	,641	,315	,959	.			
	N	18	8	11	15	0	26		
J	Korrelation nach Pearson	-,055	-,073	,363 [*]	,019	,116	-,024	1	
	Signifikanz (2-seitig)	,727	,748	,014	,935	,414	,938		
	N	43	22	45	20	52	13	65	
P	Korrelation nach Pearson	,060	-,330	-,272	,321	-,357	,418	. ^a	1
	Signifikanz (2-seitig)	,771	,385	,308	,181	,102	,155	.	
	N	26	9	16	19	22	13	0	35

a. Kann nicht berechnet werden, da mindestens eine der Variablen konstant ist.

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Bezüglich der Zusammenhänge der Merkmale des MBTI fallen zwei signifikante Korrelationen auf. Die Präferenz „Thinking“ (T) korreliert signifikant mit der Präferenz „Extraversion“ (E). Das heißt, dass extravertierte Menschen Entscheidungen durch analytisches Vorgehen treffen und Probleme anhand von Logik lösen. So wird die T-Präferenz vom Kirby und Myers beschrieben.¹⁷¹ Durch die Beschreibung der Präferenzen lässt sich diese Korrelation nicht abschließend klären, denn es gibt keine ersichtlichen Überschneidungen zwischen diesen beiden Präferenzen. Ein anderer Erklärungsansatz liegt in der Verteilung der Häufigkeit der Präferenzen. Die Präferenzen E, S, T und J treten insgesamt häufiger auf, sie sind allgemein wahrscheinlicher. Dieser Aspekt wird auch schon in der Berechnung der Rohwerte berücksichtigt. Es ist also möglich, dass diese Korrelation allein deswegen entsteht, dass die Wahrscheinlichkeit des Auftretens beider Merkmale höher ist und die vorliegende Stichprobe damit diese Korrelation aufweist. Dafür spricht auch der geringe Wert der Korrelation.

Die zweite signifikante Korrelation ist keine Überraschung. Das Merkmal „Judging“ (J) korreliert signifikant in mäßiger Höhe mit der Präferenz „Sensing“ (S). Kirby und Myers nennen im Zusammenhang mit S Eigenschaften, die die Informationsaufnahme betreffen. S-Typen nehmen gerne Informationen auf, die realistisch und greifbar sind und konzentrieren sich auf das, was wirklich passiert. Sie sind faktenbezogen und konkret und kommen durch sorgfältige und gründliche Erwägungen zu Schlussfolgerungen. Dieser Aspekt der Sorgfältigkeit und des Strukturierens wird auch in der J-Präferenz ein wenig aufgegriffen. J-Typen sind planerisch veranlagt und organisieren ihr Leben. Sie sind systematisch und gehen methodisch vor. Dieser Auszug aus den Eigenschaften, die von Kirby und Myers für Personen mit J-Präferenz genannt werden, verdeutlicht, dass die beiden Präferenzen durchaus eine gemeinsame Schnittmenge von Eigenschaften besitzen, die freilich nicht sehr groß ist, aber für die Erklärung der geringen Korrelation ausreicht.

6.3 Korrelativer Vergleich zwischen den Verfahren

In diesem Abschnitt werden die Korrelationen zwischen den Verfahren betrachtet. Damit sollen Zusammenhänge zwischen den Verfahren aufgezeigt und erklärt werden. Zunächst wird jeweils der Myers-Briggs Typenindikator als Typologie gegen die beiden traitorientierten Verfahren – NEO-Fünf-Faktoren-Inventar und Eysenck-Personality-Questionnaire gestellt. Es soll überprüft werden, ob es trotz der unterschiedlichen

¹⁷¹ vgl. Kirby/ Myers, 2001, S. 10

Konzeptionen der Verfahren, Überschneidungen der Merkmale gibt. Im letzten Schritt werden der NEO-FFI und der EPQ-RK verglichen.

Eysenck Personality Questionnaire und Myers-Briggs Typenindikator

In Tabelle 13 ist das Ergebnis der Korrelationsanalyse zwischen dem EPQ-RK und dem MBTI abgebildet.

Tab. 13: Die Korrelationen zwischen dem EPQ-RK und dem MBTI

	EPQ N	EPQ P	EPQ E	E	I	S	N	T	F	J	P
EPQ N Korrelation nach Pearson Signifikanz (2-seitig) N	1 100										
EPQ P Korrelation nach Pearson Signifikanz (2-seitig) N	,054 100	1 100									
EPQ E Korrelation nach Pearson Signifikanz (2-seitig) N	,188 100	,129 100	1 100								
E Korrelation nach Pearson Signifikanz (2-seitig) N	,215 69	,020 69	,551** 69	1 69							
I Korrelation nach Pearson Signifikanz (2-seitig) N	,368* 31	,225 31	,598** 31	. ^a 0	1 31						
S Korrelation nach Pearson Signifikanz (2-seitig) N	,074 61	,308* 61	-,205 61	,002 42	,150 19	1 61					

Fortsetzung der Tabelle auf der nächsten Seite

	EPQ N	EPQ P	EPQ E	E	I	S	N	T	F	J	P
N Korrelation nach Pearson	,122	,260	,084	,077	,358	. ^a	1				
Signifikanz (2-seitig)	,460	,110	,612	,704	,254	.					
N	39	39	39	27	12	0	39				
T Korrelation nach Pearson	,134	,080	,097	,298 [*]	,084	,084	,241	1			
Signifikanz (2-seitig)	,254	,497	,409	,034	,703	,562	,257				
N	74	74	74	51	23	50	24	74			
F Korrelation nach Pearson	,084	,191	,296	,118	,196	,334	,015	. ^a	1		
Signifikanz (2-seitig)	,684	,351	,142	,641	,641	,315	,959	.			
N	26	26	26	18	8	11	15	0	26		
J Korrelation nach Pearson	,279 [*]	,169	,257 [*]	,055	,073	,363 [*]	,019	,116	,024	1	
Signifikanz (2-seitig)	,025	,178	,039	,727	,748	,014	,935	,414	,938		
N	65	65	65	43	22	45	20	52	13	65	
P Korrelation nach Pearson	,061	,402 [*]	,114	,060	,330	,272	,321	,357	,418	. ^a	1
Signifikanz (2-seitig)	,728	,017	,513	,771	,385	,308	,181	,102	,155	.	
N	35	35	35	26	9	16	19	22	13	0	35

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

a. Kann nicht berechnet werden, da mindestens eine der Variablen konstant ist.

Aus den empirischen Belegen (Abschnitt 4.2.) war zu erwarten, dass die EI-Skala des MBTI hoch mit der Skala Extraversion des EPQ korreliert. Das Ergebnis der Studie von Francis und Jones kann bestätigt werden. Das Merkmal Extraversion des EPQ-RK korreliert hoch signifikant in positiver Richtung mit dem Merkmal Extraversion des MBTI und hoch signifikant in negativer Richtung mit dem Merkmal Introversion des MBTI. Diese Korrelationen implizieren, dass beide unabhängige Methoden gleichartige psychologische Konstrukte messen.

Weiterhin korreliert die Präferenz Introversion des MBTI signifikant, aber nicht stark positiv mit dem Merkmal Neurotizismus des EPQ-RK. Das deutet an, dass die Konzeptualisierung des MBTI von I und E die Extravertierten als stabiler und die Introvertierten als weniger stabil konstruiert.

Das Merkmal „Sensing“ des MBTI weist einen signifikanten negativen Zusammenhang mit dem Merkmal Psychotizismus des EPQ-RK auf. Das bedeutet, dass S-Typen

weniger psychotizistisch sind. Das MBTI-Konzept charakterisiert S-Typen als Menschen, die das Empfinden bevorzugen und gerne Informationen aufnehmen, die realistisch und greifbar sind, und sich auf das konzentrieren, was wirklich passiert. Das Merkmal Psychotizismus wird u.a. mit den Adjektiven kalt, unpersönlich, unfreundlich, nicht vertrauensvoll, seltsam, skurril beschrieben. Menschen, die zum Psychotizismus neigen haben oft paranoide Vorstellungen, dass Menschen gegen sie sind und arbeiten.¹⁷² Die Präferenz S beinhaltet Eigenschaften, die dem Merkmal Psychotizismus entgegenstehen. Damit lässt sich die gegenläufige Korrelation zwischen den beiden Skalen erklären.

Aus den Ergebnissen der vorangegangenen Studien wäre zu erwarten gewesen, dass Korrelationen auch zwischen der „Thinking“-„Feeling“-Skala des MBTI und dem Merkmal Neurotizismus des EPQ auftreten. Diese Ergebnisse können mit dieser Stichprobe nicht bestätigt werden.

In der vorliegenden Stichprobe treten aber weiterhin signifikante Korrelationen zwischen der Präferenz „Judging“ (J) des MBTI und den Merkmalen Neurotizismus und Extraversion des EPQ auf. Beide Zusammenhänge sind negativer Natur. Das bedeutet, dass Menschen mit J-Präferenz, wahrscheinlich weniger neurotizistisch und extravertiert veranlagt sind. Während der negative Zusammenhang zwischen Extraversion und J auch von Leslie und Jones (2000) aufgezeigt wurde, scheint der Zusammenhang zwischen Neurotizismus und J wiederum ein Spezifikum der vorliegenden Stichprobe zu sein. In der Konzeption der beiden Merkmale kann dafür keine ausreichende Erklärung gefunden werden.

Ein weiterer für diese Stichprobe spezifischer Zusammenhang konnte für die „Perceiving“-Präferenz und Psychotizismus gefunden werden. Demnach neigen P-Typen eher zum Psychotizismus als J-Typen. Auch hier ist es schwierig den Zusammenhang über die Konstruktion der Skalen zu erklären. Der MBTI scheint mit seiner JP-Skala jedoch auch Facetten des Psychotizismus-Konstrukts zu messen.

Zusammenfassend ist festzustellen, dass die Ergebnisse von Leslie und Jones (2000) grundsätzlich repliziert werden konnten. Außerdem konnten weitere – wahrscheinlich für die vorliegende Stichprobe spezifische – signifikante Zusammenhänge gefunden werden.

¹⁷² vgl. Francis/ Jones, 2000, S. 377 ff.

NEO-Fünf-Faktoren Inventar und Myers-Briggs Typenindikator

Die Zusammenhänge zwischen dem NEO-Fünf-Faktoren Inventar (NEO-FFI) und dem Myers-Briggs Typenindikator sind in der folgenden Tabelle dargestellt.

Tab. 14: Die Korrelationen zwischen dem NEO-FFI und dem MBTI

	NEO N	NEO E	NEO O	NEO V	NEO G	E	I	S	N	T	F	J	P
NEO N Korr. n. Pearson Signifikanz (2-seitig) N	1 100												
NEO E Korrelation nach Pearson Signifikanz (2-seitig) N	,478** ,000 100	1 100											
NEO O Korrelation nach Pearson Signifikanz (2-seitig) N	-,053 ,602 100	,114 ,260 100	1 100										
NEO V Korrelation nach Pearson Signifikanz (2-seitig) N	-,200* ,046 100	,233* ,020 100	,027 ,793 100	1 100									
NEO G Korrelation nach Pearson Signifikanz (2-seitig) N	-,129 ,202 100	,001 ,990 100	,135 ,181 100	,275** ,006 100	1 100								
E Korrelation nach Pearson Signifikanz (2-seitig) N	-,151 ,216 69	,394** ,001 69	,080 ,511 69	,014 ,907 69	-,185 ,128 69	1 69							
I Korrelation nach Pearson Signifikanz (2-seitig) N	,415* ,020 31	-,684** ,000 31	,353 ,051 31	-,262 ,155 31	-,106 ,571 31	. ^a . 0	1 31						

Fortsetzung der Tabelle auf der nächsten Seite

		NEO N	NEO E	NEO O	NEO V	NEO G	E	I	S	N	T	F	J	P
S	Korrelation nach Pearson	,095	-,127	,245	,138	,177	,002	-,150	1					
	Signifikanz (2-seitig)	,468	,330	,057	,288	,172	,989	,540						
	N	61	61	61	61	61	42	19	61					
N	Korrelation nach Pearson	,014	-,039	,385 [*]	,072	-,293	,077	-,358	. ^a	1				
	Signifikanz (2-seitig)	,932	,815	,015	,665	,070	,704	,254	.					
	N	39	39	39	39	39	27	12	0	39				
T	Korrelation nach Pearson	-,176	-,005	,194	,266 [*]	,026	,298 [*]	-,084	,084	,241				
	Signifikanz (2-seitig)	,135	,969	,097	,022	,825	,034	,703	,562	,257				
	N	74	74	74	74	74	51	23	50	24				
F	Korrelation nach Pearson	-,165	-,085	,041	,313	,070	-,118	,196	,334	,015	. ^a	1		
	Signifikanz (2-seitig)	,421	,679	,843	,120	,736	,641	,641	,315	,959	.			
	N	26	26	26	26	26	18	8	11	15	0	26		
J	Korrelation nach Pearson	-,193	-,187	,195	,306 [*]	,515 ^{**}	-,055	-,073	,363 [*]	,019	,116	,024	1	
	Signifikanz (2-seitig)	,123	,137	,120	,013	,000	,727	,748	,014	,935	,414	,938		
	N	65	65	65	65	65	43	22	45	20	52	13	65	
P	Korrelation nach Pearson	,057	-,094	,312	,004	,688 ^{**}	,060	-,330	,272	,321	-,357	,418	. ^a	1
	Signifikanz (2-seitig)	,747	,591	,068	,983	,000	,771	,385	,308	,181	,102	,155	.	
	N	35	35	35	35	35	26	9	16	19	22	13	0	35

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

* . Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

a. Kann nicht berechnet werden, da mindestens eine der Variablen konstant ist.

Aus den empirischen Belegen aus Abschnitt 4.1 ist bei dem Vergleich zwischen NEO-FFI und MBTI ein signifikanter Zusammenhang zwischen der Extraversion (E) – Introversion (I) – Skala des MBTI und dem Merkmal Extraversion des NEO-FFI zu erwarten. Auch in der vorliegenden Stichprobe tritt dieser Zusammenhang auf. Dabei korreliert die Präferenz E des MBTI hoch signifikant positiv und I hoch signifikant negativ mit dem Merkmal Extraversion des NEO-FFI. Auch die Stärke des Zusammenhangs ist beachtlich. Damit kann gezeigt werden, dass beide Verfahren ein gleichartiges psychologisches Konstrukt messen.

Eine weitere signifikante Korrelation tritt zwischen I des MBTI und dem Merkmal Neurotizismus des NEO-FFI auf. Der MBTI scheint Introvertierte als weniger stabil zu konstruieren. Dieser Zusammenhang wurde auch beim vorhergehenden Vergleich zwischen MBTI und EPQ-RK festgestellt.

Zwischen der Präferenz „Intuition“ (N) des MBTI und dem Merkmal „Offenheit für neue Erfahrung“ des NEO-FFI besteht ebenfalls eine signifikante Korrelation, die einen positiven Zusammenhang zwischen beiden Skalen darstellt. Dies bedeutet, dass der MBTI die intuitiven Typen als offener für neue Erfahrungen charakterisiert. Dies wird durch die Beschreibung der N-Präferenz unterstützt. Menschen, die die Intuition bevorzugen sind u.a. besonders darauf eingestellt, neue Möglichkeiten zu erkennen, sind voller Vorstellungskraft und kreativ.¹⁷³ Probanden mit hohen Werten bezüglich „Offenheit für neue Erfahrungen“ bevorzugen Abwechslung, sind kreativ, phantasievoll und wissbegierig.¹⁷⁴ Damit lassen sich eindeutige Überschneidungen zwischen den beiden Konstrukten finden und die Korrelationen erklären.

Bezüglich des Merkmals Verträglichkeit des NEO-FFI und der „Thinking“-Präferenz lässt sich ein geringer aber signifikanter negativer Zusammenhang feststellen. Folgerichtig konstruiert der MBTI T-Typen, die sehr analytisch sind, bei ihren Entscheidungen das Denken bevorzugen und die logischen Konsequenzen ihrer Entscheidung berücksichtigen als weniger verträglich. Auch hier überschneiden sich beide Konstrukte wieder einmal. T-Typen, die ihre Entscheidungen aufgrund logischer Überlegungen treffen, blenden die Interessen und Bedürfnisse ihrer Mitmenschen in diesen Entscheidungssituationen eher aus, während Menschen mit hohen Werten in der Skala Verträglichkeit sehr rücksichtsvoll und mitfühlend sind. Damit erklärt sich die negative Korrelation zwischen der T-Präferenz des MBTI und dem Merkmal Verträglichkeit des NEO-FFI. Es ist verwunderlich, dass keine signifikante positive Korrelation zwischen der S-Präferenz und der Verträglichkeit festzustellen ist, da hier ebenfalls eine eindeutige Überschneidung der zugrundeliegenden psychologischen Konstrukte nachgewiesen werden kann.

Der signifikante positive Zusammenhang zwischen der Präferenz „Judging“ (MBTI) und dem Merkmal Verträglichkeit (NEO-FFI) ist dagegen nicht so offensichtlich. Er kann damit erklärt werden, dass Menschen, die ihr Leben sehr organisieren und planerisch veranlagt sind, die ihre Dinge gerne geregelt sehen, wohl eher berechenbarer sind, als Menschen, die auf einer sehr spontanen und flexiblen Weise leben. Der MBTI konstruiert die J-Typen damit als verträglicher als die P-Typen. Dieser Zusammenhang konnte auch schon in anderen Studien nachgewiesen werden.

¹⁷³ vgl. Kirby/ Myers, 2000, S.9

¹⁷⁴ vgl. Borkenau/ Ostendorf, 1993, S. 5

Auch nachgewiesen werden konnte der Zusammenhang der „Judging“(J)-„Perceiving“(P)-Skala mit dem Merkmal der Gewissenhaftigkeit. Die J-Typen wurden bereits als sehr strukturiert beschrieben, während die P-Typen eher als spontan bezeichnet werden und detaillierte Pläne als Einschränkung ansehen. Das Merkmal Gewissenhaftigkeit des NEO-FFI wird dagegen u.a. beschrieben mit diszipliniert, pünktlich, ordentlich und systematisch. Die Überschneidungen der beiden Konstrukte sind eindeutig erkennbar und damit die Zusammenhänge erklärbar.

Auch beim Vergleich dieser beiden Verfahren konnten die Ergebnisse vorangegangener Studien weitgehend repliziert werden.

Eysenck Personality Questionnaire (RK) und NEO-Fünf-Faktoren Inventar

Für den Vergleich zwischen dem revidierten Eysenck Personality Questionnaire in der Kurzform (EPQ-RK) und dem NEO-Fünf-Faktoren Inventar (NEO-FFI) konnten keine empirischen Belege gefunden werden.

Bei beiden Verfahren handelt es sich um traitorientierte Verfahren, die jedoch unterschiedliche Persönlichkeitsmodelle zur Grundlage haben. Während Eysenck lediglich drei Faktoren zur Beschreibung der Persönlichkeit nutzt, präsentiert der NEO-FFI das Fünf-Faktoren-Modell, welches fünf grundlegende Merkmale zur Beschreibung der Persönlichkeit konstatiert.

Nachfolgend sind die Korrelationen zwischen den beiden traitorientierten Verfahren EPQ-RK und NEO-FFI tabellarisch aufbereitet.

Tab. 15: Die Korrelationen des EPQ-RK und des NEO-FFI

		EPQ_E	EPQ_N	EPQ_P	NEO_N	NEO_E	NEO_O	NEO_V	NEO_G
EPQ_E	Korrelation nach Pearson	1							
	Signifikanz (2-seitig)								
	N	100							
EPQ_N	Korrelation nach Pearson	-,188	1						
	Signifikanz (2-seitig)	,061							
	N	100	100						
EPQ_P	Korrelation nach Pearson	,129	,054	1					
	Signifikanz (2-seitig)	,201	,596						
	N	100	100	100					
NEO_N	Korrelation nach Pearson	-,299**	,665**	,093	1				
	Signifikanz (2-seitig)	,003	,000	,359					
	N	100	100	100	100				
NEO_E	Korrelation nach Pearson	,766**	-,288**	-,071	-,478**	1			
	Signifikanz (2-seitig)	,000	,004	,483	,000				
	N	100	100	100	100	100			
NEO_O	Korrelation nach Pearson	,203*	,048	,027	-,053	,114	1		
	Signifikanz (2-seitig)	,043	,638	,786	,602	,260			
	N	100	100	100	100	100	100		
NEO_V	Korrelation nach Pearson	,049	-,177	-,208*	-,200*	,233*	-,027	1	
	Signifikanz (2-seitig)	,629	,078	,038	,046	,020	,793		
	N	100	100	100	100	100	100	100	
NEO_G	Korrelation nach Pearson	-,110	-,102	-,423**	-,129	,001	-,135	,275**	1
	Signifikanz (2-seitig)	,276	,310	,000	,202	,990	,181	,006	
	N	100	100	100	100	100	100	100	100

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

* . Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Die hoch signifikanten positiven Zusammenhänge zwischen den jeweils gleichnamigen Merkmalen Neurotizismus und Extraversion der beiden Verfahren sind zu erwarten gewesen. Diese beiden Merkmale korrelieren auch untereinander, d.h. es korreliert jeweils das Merkmal Neurotizismus mit dem Merkmal Extraversion in positive Richtung. Betrachtet man jeweils die Interkorrelationen in den beiden Verfahren ist auch das zu erwarten gewesen. Diese Ergebnisse implizieren, dass beide unabhängige Instrumente die gleichen psychologischen Konstrukte messen. Die Verfahren definieren die betreffenden Merkmale über die gleichen Eigenschaften.

Das Merkmal Psychotizismus wird vom NEO-FFI als solches nicht erfasst. Es treten jedoch signifikante Korrelationen zwischen Psychotizismus und den Skalen Verträglichkeit und Gewissenhaftigkeit des NEO-FFI auf. Betrachtet man hier wieder die Eigenschaften, die die Skalen charakterisieren, kann man auch hier tatsächlich Überschneidungen feststellen. Nach Eysenck kann man Psychotizismus mit den Adjektiven kalt, unpersönlich, feindselig, seltsam, unsympathisch und skurril u.a. beschreiben. Der negative Zusammenhang zum Merkmal Verträglichkeit lässt sich damit eindeutig aufklären. Das Menschen die hohe Werte beim Merkmal Psychotizismus aufweisen, auch weniger gewissenhaft sind, deutet darauf hin, dass der NEO-FFI mit dem Merkmal Gewissenhaftigkeit auch Aspekte des Psychotizismus-Konstruktes nach Eysenck erfasst.

6.4 Zusammenfassende Ergebnisse

Betrachtet man die Ergebnisse der Vergleiche insgesamt, kann man feststellen, dass sich alle drei Verfahren in vielen Facetten überschneiden. Die drei Verfahren messen auf unterschiedliche Art und Weise grundsätzlich dieselben psychologischen Konstrukte, um die Persönlichkeit eines Menschen näher zu beschreiben. Sie unterscheiden sich nur mehr in den Persönlichkeitsmodellen, auf deren Grundlage sie entwickelt worden sind.

Weiterhin ist festzustellen, dass die Zusammenhänge zwischen den beiden traitorientierten Verfahren EPQ-RK und NEO-FFI deutlich einfacher nachvollziehbar und erklärbar sind, als die Zusammenhänge zwischen der Typologie MBTI und den anderen beiden Verfahren. Dies kann der Konstruktion einer Typologie im Allgemeinen oder speziell der des MBTI geschuldet sein.

Dass der MBTI ein sehr umstrittenes Verfahren darstellt ist nicht verwunderlich. Vergleicht man die die MBTI-Dichotomien mit den Faktoren des NEO-FFI und des EPQ-RK findet man Zusammenhänge, die teilweise wenig nachvollziehbar sind. Andererseits überschneiden sich die Konzeptionen der Präferenzen des MBTI und der Merkmale des NEO-FFI und des EPQ-RK in einigen Fällen doch extrem. McCrae und Costa (1989) plädieren in ihrer Studie für eine radikale Neuinterpretation des MBTI aus der Sicht des Fünf-Faktoren Modells. Es ist in mehreren Untersuchungen nachgewiesen worden, dass der MBTI die Typentheorie von Jung nicht sauber abbildet. Das Konzept stellt sich bei näherer Betrachtung als unstrukturiert und verzerrt heraus. Außerdem ist noch kein Beweis gefunden worden, dass die Präferenzen klare Dichotomien formen. McCrae und Costa (1989) sehen die Konstruktvalidität des MBTI als stark gefährdet. Der MBTI scheint kein Instrument zu sein, dass tatsächlich die Typen nach Jung misst und entbehrt damit eigentlich seiner theoretischen Grundlage.¹⁷⁵

Trotz der heftigen Kritiken ist der MBTI immer noch eines der am häufigsten eingesetzten Verfahren wenn es um Personal- und Teamentwicklung geht.¹⁷⁶ Die Popularität des MBTI resultiert aus der einfachen Anwendung des Instrumentes. Hat man nicht den Anspruch die Typen nach Jung messen zu wollen, spricht nichts gegen den Einsatz des MBTI z.B. in der Teamentwicklung oder beim Coaching. Er sollte jedoch keinesfalls als Selektionsinstrument genutzt werden.

Die Validität der Verfahren (siehe Kapitel 3) spricht mehr für den Einsatz der beiden traitorientierten Verfahren. Der EPQ-RK ist für den Einsatz im wirtschaftlichen Kontext wahrscheinlich weniger geeignet als der NEO-FFI, da eine Beziehung der Items zu berufsrelevanten Fragestellungen – vor allem beim Merkmal Psychotizismus – oft schwierig herzustellen ist. Auch hier wird nochmals darauf verwiesen, dass auch diese Verfahren zwar die prognostische Validität eines eignungsdiagnostischen Gesamtkonzeptes deutlich erhöhen können, aber keinesfalls als alleinige Selektionsinstrumente genutzt werden sollten.¹⁷⁷

¹⁷⁵ vgl. McCrae/ Costa, 1989, S. 17 ff.

¹⁷⁶ vgl. ebenda

¹⁷⁷ vgl. Hossiep/ Mühlhaus, 2005, S. 9

6.5 Methodische Anmerkungen und Einschränkungen der vorliegenden Untersuchung

Ohne den Wert der vorliegenden Untersuchung zu schmälern, scheinen einige kritische Anmerkungen zur Vorgehensweise und Durchführung der Untersuchung angebracht. Diese sollen zur besseren Einordnung der Ergebnisse dienen.

Im Sinne der Zielsetzung, empirische Zusammenhänge zwischen den drei ausgewählten Verfahren, zu überprüfen, reicht der Stichprobenumfang nicht aus, um allgemeingültige Aussagen zu treffen. Da die Teilnehmer fast ausschließlich Studenten der Hochschule Mittweida waren, können lediglich Aussagen bezüglich dieser Personengruppe getroffen werden. Es ist durchaus vorstellbar, dass Studenten die Intention der gestellten Fragen eher erkennen als Absolventen eines anderen Bildungsweges und die Ergebnisse damit unbewusst beeinflussen.

Bei der Ergebnisinterpretation muss weiterhin berücksichtigt werden, dass die Teilnahme an der Untersuchung prinzipiell freiwillig war. Bezüglich der Ergebnisse wurde den Teilnehmern Anonymität und Datenschutz zugesichert. Sie mussten mit keinerlei Konsequenzen rechnen und hatten die Möglichkeit, eine Rückmeldung ihrer Ergebnisse zu erhalten. Diese Rahmenbedingungen könnten das Antwortverhalten der Teilnehmer ebenfalls beeinflusst haben.

Die Datenanalyse und -auswertung aufgrund von Korrelationsbestimmungen stellt einen geeigneten Ansatz dar und ist für die Beantwortung der zentralen Fragestellungen dieser Arbeit ausreichend. Vor dem Hintergrund zahlreicher Untersuchungen, die Geschlechterunterschiede bei Persönlichkeitsverfahren belegen, wäre jedoch eine ausführlichere geschlechtervergleichende Datenanalyse sinnvoll. Um diese durchführen zu können, ist anzuraten, den Stichprobenumfang für eine weitere Untersuchung zu erhöhen und dabei auf eine Gleichverteilung von weiblichen und männlichen Probanden zu achten.

7 Perspektiven

Persönlichkeitsverfahren im wirtschaftlichen Kontext sind nach wie vor nicht unumstritten. Die Diskussionen um Sinn und Nutzen dieser wird auch in den nächsten Jahren fortgeführt werden. Dennoch sind die meisten Forscher, die dieses Thema untersuchen, sich weitgehend einig, dass gut konstruierte Persönlichkeitsdimensionen positionsübergreifend als valide Prädiktoren für den Berufserfolg dienen können.¹⁷⁸

Persönlichkeitsverfahren im Personalmanagement erfahren eine steigende Akzeptanz, wobei sowohl die Anwender als auch die Teilnehmer deutlich sensibler und kritischer mit den Verfahren umgehen. Es passiert immer weniger, dass klinisch-psychologische Verfahren unreflektiert für Auswahlentscheidungen eingesetzt werden. Die Zukunft der psychologischen Eignungsdiagnostik im wirtschaftlichen Kontext liegt in den berufsbezogenen Verfahren. Das wachsende Angebot an diesen Verfahren wird dazu führen, dass sie zunehmend mehr in die Personalarbeit integriert werden. Zu berücksichtigen ist dabei auch die Akzeptanz bei den Bewerbern. Werden die Testverfahren als nicht sinnvoll und nützlich erlebt, werden die Bewerber mit Blick auf das Image des Unternehmens eher zurückhaltender sein. Dies bedeutet für das Unternehmen, dass sich potentielle Mitarbeiter im Vorfeld selbst selektieren. Berücksichtigt man auch die demographische Entwicklung in Deutschland, kann sich dieses Problem extrem verschärfen, denn den Unternehmen werden in Zukunft immer weniger qualifizierte (potentielle) Mitarbeiter zur Verfügung stehen. Folgerichtig sollte die Frage der Akzeptanz von Testverfahren bei den Bewerbern und Mitarbeitern nicht unterschätzt werden.

Neben umfassenden berufsbezogenen Instrumenten wie z.B. dem Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (Hossiep und Paschen, 1998) wird die praktische Personalarbeit in der Zukunft wahrscheinlich auch nach spezifischen Instrumenten verlangen, die einzelne, für eine Entscheidung wichtige Subkonstrukte erfassen (vgl. Leistungsmotivationsinventar von Schuler und Prochaska). Dazu gehört u.a. die Kundenorientierung oder auch die Integrität einer Person. Während in Deutschland bisher nur wenig über die Erfassung der Integrität geforscht wird, wird diese Thematik im US-amerikanischen Raum schon seit vielen Jahren verfolgt. Hierzulande boomt der Markt der Beratungsgesellschaften, die diese Marktlücke entdeckt haben und diesbezüglich selbstentwickelte Instrumente anbieten – bei denen aber die Seriosität der Konstruktion häufig nicht nachvollzogen werden kann. Ein

¹⁷⁸ vgl. Hogan/ Hogan/ Roberts, 1996, S. 469 ff.

stärkeres Engagement von Seiten der wissenschaftlichen Psychologie und der anwendungsbezogenen wissenschaftlichen Forschung scheint diesbezüglich erforderlich.¹⁷⁹

Die Diagnostik von Werthaltungen stellt ein weiteres Anwendungsgebiet mit Bedarf an neuen Instrumenten dar. Hinsichtlich von Positionen, bei denen die Bewerber vermutlich ohnehin schon über die notwendigen fachlichen und sozialen Kompetenzen verfügen oder die Vorauswahl unter Berücksichtigung dieser Kriterien schon abgeschlossen ist, steht vor allem die Einstellung und Passung zu Kultur und Werten des Unternehmens im Vordergrund. Auch dieser Prozess könnte durch Fragebogenverfahren sinnvoll unterstützt werden. Werthaltungen hängen nicht nur mit Persönlichkeitsfaktoren zusammen, sondern determinieren auch das Verhalten in Organisationen in den verschiedensten Situationen.¹⁸⁰

Die Frage ob Persönlichkeitsverfahren einen positiven Beitrag zur praktischen Personalarbeit leisten können ist grundsätzlich zu bejahen. Dabei ist es nicht nur eine Frage der wissenschaftlich nachgewiesenen Nützlichkeit. Vielmehr ist es wichtig, dass diese Verfahren ihren Wert nur in der Hand eines sachverständigen Diagnostikers tatsächlich entfalten können. Obwohl die Sensibilität dahingehend schon gestiegen ist, wird dieser Aspekt dennoch, aufgrund des zunehmenden Drucks auf viele Personalbereiche, immer wieder unberücksichtigt gelassen.

Das Personal stellt die wichtigste Ressource jedes Unternehmens dar. Wer mit dieser wichtigsten Ressource zu tun hat, sollte sich den Freiraum erkämpfen, richtige und zukunftsweisende Maßnahmen ergreifen zu können. Dazu gehört eine qualitativ hochwertige Personalarbeit in allen Bereichen, da sich der Wettbewerb der Zukunft genau auf diesem Gebiet entscheiden wird. Das Unternehmen, welches das Potential seiner Mitarbeiter kennt und dementsprechend einzusetzen weiß, wird sich in der Zukunft entscheidende Wettbewerbsvorteile sichern können.

¹⁷⁹ vgl. Hossiep/ Paschen/ Mühlhaus, 2000, S. 262

¹⁸⁰ vgl. ebenda

Anhang

Tabellen zum Test auf Normalverteilung

Myers-Briggs Typenindikator

Kolmogorov-Smirnov-Anpassungstest									
		MBTI_E	MBTI_I	MBTI_S	MBTI_N	MBTI_T	MBTI_F	MBTI_J	MBTI_P
N		69	31	61	39	74	26	65	35
Parameter der	Mittelwert	29,8116	26,8710	16,4918	14,7949	24,8784	13,4615	26,9077	20,9429
Normalverteilung ^{a,b}	Standardabweichung	16,85690	16,55645	11,16038	10,78786	16,62145	7,74199	14,26723	14,32797
Extremste Differenzen	Absolut	,085	,109	,121	,176	,115	,144	,069	,147
	Positiv	,073	,109	,121	,176	,115	,144	,069	,147
	Negativ	-,085	-,091	-,083	-,137	-,075	-,117	-,054	-,096
Kolmogorov-Smirnov-Z		,703	,604	,946	1,099	,987	,735	,560	,867
Asymptotische Signifikanz (2-seitig)		,706	,858	,332	,179	,285	,652	,912	,440

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

NEO-Fünf-Faktoren-Inventar

Kolmogorov-Smirnov-Anpassungstest						
		Neurotizismus	Extraversion	Offenheit für Erfahrung	Verträglichkeit	Gewissenhaftigkeit
N		100	100	100	100	100
Parameter der	Mittelwert	1,44840	2,5288	2,32467	2,6017	2,8269
Normalverteilung ^{a,b}	Standardabweichung	,551285	,48716	,480254	,45318	,52124
Extremste Differenzen	Absolut	,095	,142	,076	,068	,062
	Positiv	,095	,069	,076	,049	,060
	Negativ	-,047	-,142	-,063	-,068	-,062
Kolmogorov-Smirnov-Z		,950	1,419	,763	,683	,617
Asymptotische Signifikanz (2-seitig)		,327	,036	,605	,740	,841

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

Eysenck Personality Questionnaire**Kolmogorov-Smirnov-Anpassungstest**

		N	P	E
N		100	100	100
Parameter der Normalverteilung ^{a,b}	Mittelwert	3,6200	2,9000	8,1900
	Standardabweichung	2,54963	1,71447	3,46904
Extremste Differenzen	Absolut	,151	,180	,159
	Positiv	,151	,180	,136
	Negativ	-,082	-,094	-,159
Kolmogorov-Smirnov-Z		1,508	1,802	1,591
Asymptotische Signifikanz (2-seitig)		,021	,003	,013

a. Die zu testende Verteilung ist eine Normalverteilung.

b. Aus den Daten berechnet.

Anhang

Tabellen zu den Interkorrelationen in den Verfahren

Myers-Briggs Typenindikator

		Korrelationen							
		MBTI_E	MBTI_I	MBTI_S	MBTI_N	MBTI_T	MBTI_F	MBTI_J	MBTI_P
MBTI_E	Korrelation nach Pearson	1 ^a		,002	,077	,298	-,118	-,055	,060
	Signifikanz (2-seitig)		.	,989	,704	,034	,641	,727	,771
	N	69	0	42	27	51	18	43	26
MBTI_I	Korrelation nach Pearson	^a	1	-,150	-,358	-,084	,196	-,073	-,330
	Signifikanz (2-seitig)	.		,540	,254	,703	,641	,748	,385
	N	0	31	19	12	23	8	22	9
MBTI_S	Korrelation nach Pearson	,002	-,150	1 ^a		,084	,334	,363	-,272
	Signifikanz (2-seitig)	,989	,540	.		,562	,315	,014	,308
	N	42	19	61	0	50	11	45	16
MBTI_N	Korrelation nach Pearson	,077	-,358 ^a		1	-,241	-,015	,019	,321
	Signifikanz (2-seitig)	,704	,254	.		,257	,959	,935	,181
	N	27	12	0	39	24	15	20	19
MBTI_T	Korrelation nach Pearson	,298	-,084	,084	-,241	1 ^a		,116	-,357
	Signifikanz (2-seitig)	,034	,703	,562	,257	.		,414	,102
	N	51	23	50	24	74	0	52	22
MBTI_F	Korrelation nach Pearson	-,118	,196	,334	-,015 ^a		1	-,024	,418
	Signifikanz (2-seitig)	,641	,641	,315	,959	.		,938	,155
	N	18	8	11	15	0	26	13	13
MBTI_J	Korrelation nach Pearson	-,055	-,073	,363	,019	,116	-,024	1 ^a	.
	Signifikanz (2-seitig)	,727	,748	,014	,935	,414	,938	.	.
	N	43	22	45	20	52	13	65	0
MBTI_P	Korrelation nach Pearson	,060	-,330	-,272	,321	-,357	,418 ^a		1
	Signifikanz (2-seitig)	,771	,385	,308	,181	,102	,155	.	.
	N	26	9	16	19	22	13	0	35

a. Kann nicht berechnet werden, da mindestens eine der Variablen konstant ist.

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

NEO-Fünf-Faktoren-Inventar

		Korrelationen				
		a_NEO- FFI_Neurotizismus (scalar)_ROHWERT	a_NEO- FFI_Extraversion (scalar)_ROHWERT	a_NEO- FFI_Offenheit für Erfahrung (scalar)_ROHWERT	a_NEO- FFI_Verträglichkeit (scalar)_ROHWERT	a_NEO- FFI_Gewissenhaftigkeit (scalar)_ROHWERT
a_NEO-FFI_Neurotizismus (scalar)_ROHWERT	Korrelation nach Pearson	1		-,053	-,200*	-,129
	Signifikanz (2-seitig)		,000	,602	,046	,202
	N	100	100	100	100	100
a_NEO-FFI_Extraversion (scalar)_ROHWERT	Korrelation nach Pearson	-,478**	1	,114	,233*	,001
	Signifikanz (2-seitig)	,000		,260	,020	,990
	N	100	100	100	100	100
a_NEO-FFI_Offenheit für Erfahrung (scalar)_ROHWERT	Korrelation nach Pearson	-,053	,114	1	-,027	-,135
	Signifikanz (2-seitig)	,602	,260		,793	,181
	N	100	100	100	100	100
a_NEO-FFI_Verträglichkeit (scalar)_ROHWERT	Korrelation nach Pearson	-,200*	,233*	-,027	1	,275**
	Signifikanz (2-seitig)	,046	,020	,793		,006
	N	100	100	100	100	100
a_NEO- FFI_Gewissenhaftigkeit (scalar)_ROHWERT	Korrelation nach Pearson	-,129	,001	-,135	,275**	1
	Signifikanz (2-seitig)	,202	,990	,181	,006	
	N	100	100	100	100	100

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Eysenck Personality Questionnaire**Korrelationen**

		a_NEO- FFI_Neurotizismus (scalar)_ROHWERT	a_NEO- FFI_Extraversion (scalar)_ROHWERT	a_NEO- FFI_Offenheit für Erfahrung (scalar)_ROHWERT	a_NEO- FFI_Verträglichkeit (scalar)_ROHWERT	a_NEO- FFI_Gewissenhaftigkeit (scalar)_ROHWERT
a_NEO-FFI_Neurotizismus (scalar)_ROHWERT	Korrelation nach Pearson	1		-,053	-,200*	-,129
	Signifikanz (2-seitig)		,000	,602	,046	,202
	N	100	100	100	100	100
a_NEO-FFI_Extraversion (scalar)_ROHWERT	Korrelation nach Pearson	-,478**	1	,114	,233*	,001
	Signifikanz (2-seitig)	,000		,260	,020	,990
	N	100	100	100	100	100
a_NEO-FFI_Offenheit für Erfahrung (scalar)_ROHWERT	Korrelation nach Pearson	-,053	,114	1	-,027	-,135
	Signifikanz (2-seitig)	,602	,260		,793	,181
	N	100	100	100	100	100
a_NEO-FFI_Verträglichkeit (scalar)_ROHWERT	Korrelation nach Pearson	-,200*	,233*	-,027	1	,275**
	Signifikanz (2-seitig)	,046	,020	,793		,006
	N	100	100	100	100	100
a_NEO- FFI_Gewissenhaftigkeit (scalar)_ROHWERT	Korrelation nach Pearson	-,129	,001	-,135	,275**	1
	Signifikanz (2-seitig)	,202	,990	,181	,006	
	N	100	100	100	100	100

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

* . Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Tabellen zu den Korrelationen zwischen den Verfahren

Myers-Briggs-Typenindikator und NEO-Fünf-Faktoren Inventar

Korrelationen												
	a_NEO- FFI_Neurotizismus (scalar)_ROHW ERT	a_NEO- FFI_Extraversion (scalar)_ROHW ERT	a_NEO- FFI_Offenhheit für Erfahrung (scalar)_ROHW ERT	a_NEO- FFI_Verträglichkeit (scalar)_ROHW ERT	a_NEO- FFI_Gewissenhaftigkeit (scalar)_ROHW ERT	MBTI_E	MBTI_I	MBTI_S	MBTI_N	MBTI_T	MBTI_F	MBTI_J
a_NEO-FFI_Neurotizismus (scalar)_ROHWERT	1	-.478 ^{**}	-.063	-.200	-.129	-.151	-.415 ^{**}	.095	.014	-.176	-.165	-.183
N	100	100	100	100	100	69	31	61	39	74	26	65
a_NEO-FFI_Extraversion (scalar)_ROHWERT	-.478 ^{**}	1	.114	.233 ^{**}	.001	.394 ^{**}	-.684 ^{**}	-.127	-.039	-.005	-.085	-.187
N	100	100	100	100	100	69	31	61	39	74	26	65
a_NEO-FFI_Offenhheit für Erfahrung (scalar)_ROHWERT	-.053	.114	1	-.027	.135	.080	-.353	-.245	.385	-.194	-.041	-.195
N	100	100	100	100	100	69	31	61	39	74	26	65
a_NEO-FFI_Verträglichkeit (scalar)_ROHWERT	-.200	.233 ^{**}	-.027	1	.275 ^{**}	.014	-.262	.138	.072	-.266	.313	.306
N	100	100	100	100	100	69	31	61	39	74	26	65
a_NEO-FFI_Gewissenhaftigkeit (scalar)_ROHWERT	-.129	.001	-.135	.275 ^{**}	1	-.185	-.106	.177	-.293	.026	.070	.515 ^{**}
N	100	100	100	100	100	69	31	61	39	74	26	65
MBTI_E	-.151	.394 ^{**}	.080	.014	-.185	1 ^a	.014	.002	.077	.298 ^{**}	-.118	-.055
N	69	69	69	69	69	69	69	61	39	74	26	65
MBTI_I	.415 ^{**}	-.684 ^{**}	-.353	-.262	.177	-.106	1	-.150	-.358	-.084	.196	-.073
N	31	31	31	31	31	31	31	19	12	23	8	22
MBTI_S	.095	-.127	-.245	.138	.177	.014	-.150	1 ^a	.084	.334	.363	.363
N	100	100	100	100	100	100	100	100	100	.562	.315	.014
MBTI_N	.468	.330	.057	.288	.330	.989	.540	.070	.704	.034	.641	.727
N	61	61	61	61	61	61	19	61	0	50	11	45
MBTI_T	.014	-.039	.385	.072	-.293	.077	-.358 ^a	.084	1	-.241	-.015	.019
N	39	39	39	39	39	39	12	0	39	.257	.959	.935
MBTI_F	-.176	-.005	-.194	-.266	.026	.298 ^{**}	-.084	.084	-.241	1 ^a	.116	.116
N	135	.969	.097	.022	.825	.034	.703	.562	.257	.0	.414	.414
MBTI_J	-.165	-.085	-.041	.313	.515 ^{**}	-.055	-.073	.363	.019	.116	-.024	1 ^a
N	421	.679	.843	.120	.736	.641	.641	.315	.959	.0	.938	.938
MBTI_P	.057	-.084	.312	.004	-.688 ^{**}	.060	-.330	-.272	.321	-.357	.418 ^a	.418 ^a
N	35	35	35	35	35	26	9	16	19	22	13	0

**. Die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant.

*. Die Korrelation ist auf dem Niveau von 0.05 (2-seitig) signifikant.

a. Kann nicht berechnet werden, da mindestens eine der Variablen konstant ist.

Myers-Briggs Typenindikator und Eysenck Personality Questionnaire – RK

Korrelationen

	N	P	E	MBTI_E	MBTI_I	MBTI_S	MBTI_N	MBTI_T	MBTI_F	MBTI_J	MBTI_P
N											
Korrelation nach Pearson	1	,054	-,188	-,215	,368	,074	,122	-,134	-,084	-,279	,061
Signifikanz (2-seitig)		,596	,061	,076	,041	,573	,460	,254	,684	,025	,728
N	100	100	100	69	31	61	39	74	26	65	35
P											
Korrelation nach Pearson	,054	1	,129	-,020	,225	-,308	,260	,080	-,191	-,169	,402
Signifikanz (2-seitig)	,596		,201	,872	,223	,016	,110	,497	,351	,178	,017
N	100	100	100	69	31	61	39	74	26	65	35
E											
Korrelation nach Pearson	-,188	,129	1	,551	-,598	-,205	,084	,097	-,296	-,257	,114
Signifikanz (2-seitig)	,061	,201		,000	,000	,113	,612	,409	,142	,039	,513
N	100	100	100	69	31	61	39	74	26	65	35
MBTI_E											
Korrelation nach Pearson	-,215	-,020	,551	1	,1	,002	,077	,298	-,118	-,055	,060
Signifikanz (2-seitig)	,076	,872	,000			,989	,704	,034	,641	,727	,771
N	69	69	69	69	0	42	27	51	18	43	26
MBTI_I											
Korrelation nach Pearson	,368	,225	-,598		1	-,150	-,358	-,084	,196	-,073	-,330
Signifikanz (2-seitig)	,041	,223	,000			,540	,254	,703	,641	,748	,385
N	31	31	31	0	31	19	12	23	8	22	9
MBTI_S											
Korrelation nach Pearson	,074	-,308	-,205	,002	-,150	1 ^a		,084	,334	,363	-,272
Signifikanz (2-seitig)	,573	,016	,113	,989	,540			,562	,315	,014	,308
N	61	61	61	42	19	61	0	50	11	45	16
MBTI_N											
Korrelation nach Pearson	,122	,260	,084	,077	-,358		1	-,241	-,015	,019	,321
Signifikanz (2-seitig)	,460	,110	,612	,704	,254			,257	,959	,935	,181
N	39	39	39	27	12	0	39	24	15	20	19
MBTI_T											
Korrelation nach Pearson	-,134	,080	,097	,298	-,084	,084	-,241	1 ^a		,116	-,357
Signifikanz (2-seitig)	,254	,497	,409	,034	,703	,562	,257			,414	,102
N	74	74	74	51	23	50	24	74	0	52	22
MBTI_F											
Korrelation nach Pearson	-,084	-,191	-,296	-,118	,196	,334	-,015		1	-,024	,418
Signifikanz (2-seitig)	,684	,351	,142	,641	,641	,315	,959			,938	,155
N	26	26	26	18	8	11	15	0	26	13	13
MBTI_J											
Korrelation nach Pearson	-,279	-,169	-,257	-,055	-,073	,363	,019	,116	-,024	1 ^a	
Signifikanz (2-seitig)	,025	,178	,039	,727	,748	,014	,935	,414	,938		
N	65	65	65	43	22	45	20	52	13	65	0
MBTI_P											
Korrelation nach Pearson	,061	,402	,114	,060	-,330	-,272	,321	-,357	,418 ^a		1
Signifikanz (2-seitig)	,728	,017	,513	,771	,385	,308	,181	,102	,155		
N	35	35	35	26	9	16	19	22	13	0	35

*, Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

**, Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

a. Kann nicht berechnet werden, da mindestens eine der Variablen konstant ist.

NEO-Fünf-Faktoren Inventar und Eysenck Personality Questionnaire

Korrelationen									
		E	N	P	a_NEO- FFI_Neurotizismus (scalar)_ROHWERT	a_NEO- FFI_Extraversion (scalar)_ROHWERT	a_NEO- FFI_Offenheit für Erfahrung (scalar)_ROHWERT	a_NEO- FFI_Verträglichkeit (scalar)_ROHWERT	a_NEO- FFI_Gewissenhaftigkeit (scalar)_ROHWERT
E	Korrelation nach Pearson	1	-,188	,129	-,299**	,766**	,203	,049	-,110
	Signifikanz (2-seitig)		,061	,201	,003	,000	,043	,629	,276
	N	100	100	100	100	100	100	100	100
N	Korrelation nach Pearson	-,188	1	,054	,665**	-,288**	,048	-,177	-,102
	Signifikanz (2-seitig)	,061		,596	,000	,004	,638	,078	,310
	N	100	100	100	100	100	100	100	100
P	Korrelation nach Pearson	,129	,054	1	,093	-,071	,027	-,208	-,423**
	Signifikanz (2-seitig)	,201	,596		,359	,483	,786	,038	,000
	N	100	100	100	100	100	100	100	100
a_NEO-FFI_Neurotizismus (scalar)_ROHWERT	Korrelation nach Pearson	-,299**	,665**	,093	1	-,478**	-,053	-,200	-,129
	Signifikanz (2-seitig)	,003	,000	,359		,000	,602	,046	,202
	N	100	100	100	100	100	100	100	100
a_NEO-FFI_Extraversion (scalar)_ROHWERT	Korrelation nach Pearson	,766**	-,288**	-,071	-,478**	1	,114	,233	,001
	Signifikanz (2-seitig)	,000	,004	,483	,000		,260	,020	,990
	N	100	100	100	100	100	100	100	100
a_NEO-FFI_Offenheit für Erfahrung (scalar)_ROHWERT	Korrelation nach Pearson	,203	,048	,027	-,053	,114	1	-,027	-,135
	Signifikanz (2-seitig)	,043	,638	,786	,602	,260		,793	,181
	N	100	100	100	100	100	100	100	100
a_NEO-FFI_Verträglichkeit (scalar)_ROHWERT	Korrelation nach Pearson	,049	-,177	-,208	-,200	,233	-,027	1	,275**
	Signifikanz (2-seitig)	,629	,078	,038	,046	,020	,793		,006
	N	100	100	100	100	100	100	100	100
a_NEO- FFI_Gewissenhaftigkeit (scalar)_ROHWERT	Korrelation nach Pearson	-,110	-,102	-,423**	-,129	,001	-,135	,275**	1
	Signifikanz (2-seitig)	,276	,310	,000	,202	,990	,181	,006	
	N	100	100	100	100	100	100	100	100

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

* . Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Literaturverzeichnis

Bücher

Amelang, Manfred; Bartussek, Dieter; Stemmler, Gerhard; et al. (2006): Grundlagen der Differentiellen Psychologie. In: Heuer, Manfred; Rösler, Frank; Tack, Werner H. (Hrsg.): Differentielle Psychologie und Persönlichkeitsforschung. – 6., vollständig überarb. Aufl. – Stuttgart: Kohlhammer, 2006

Amelang, Manfred; Borkenau, Peter (1984): Versuch einer Differenzierung des Eigenschaftskonzepts: Aspekte intraindividuelle Variabilität und differentieller Vorhersagbarkeit. In: Amelang, Manfred; Ahrens, Hans-Joachim (Hrsg.): Brennpunkte der Persönlichkeitsforschung. – Göttingen: Hogrefe (Band 1), 1984

Amelang, Manfred; Schmidt-Atzert, Lothar (2006): Psychologische Diagnostik und Intervention. – 4. vollst. überarb. und aktualisierte Aufl. – Heidelberg: Springer Medizin, 2006

Anastasi, Anne (1985): The Use of Personality Assessment in Industry: Methodological and Interpretive Problems. In: Bernadin, H. John; Bownas, David A. (Hrsg.): Personality Assessment in Organizations. – New York: Praeger, 1985

Borkenau, Peter; Ostendorf, Fritz (1993): NEO- Fünf- Faktoren Inventar (NEO-FFI) nach Costa und Mc Crae. Handanweisung. – Göttingen: Hogrefe, 1993

Bortz, Jürgen; Döring, Nicola (2006): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. – 4., überarb. Aufl. – Heidelberg: Springer, 2006

Brox, Hans (2005): Allgemeiner Teil des Bürgerlichen Gesetzbuches. – 29. Aufl. – Köln: Carl Heymann, 2005

Comelli, Gerhard (1995): Juristische und ethische Aspekte der Eignungsdiagnostik im Managementbereich. In: Sarges, Werner (Hrsg.): Management-Diagnostik. – 2. vollständig überarb. und erweiterte Aufl.. – Göttingen: Hogrefe, 1995

Cronbach, Lee J. (1970): Essentials of Psychological Testing. – 3. Aufl. – New York: Harper&Row, 1970

Deubner, Rolf (2002): Myers-Briggs Typenindikator (MBTI). In: Brähler, Elmar; Holling, Heinz; Leutner, Detlev; et al. (Hrsg.): Brickenkamp. Handbuch psychologischer pädagogischer Tests. – 3., vollständig überarb. und erweiterte Aufl. – Göttingen: Hogrefe (Band 1), 2002

Du Bois, Philip H. (1970): A History of Psychological Testing. – Boston: Allyn and Bacon, 1970

Engelhard, Johann (1984): Entwicklungsorientierte Personalpolitik. Theoretische Grundlagen und empirische Untersuchung. – Wiesbaden: Gabler, 1984

Etzel, Steffen; Küppers, Anja (2002): Innovative Managementdiagnostik. – Göttingen: Hogrefe, 2002

Fisseni, Hermann-Josef (2004): Lehrbuch der psychologischen Diagnostik. – 3., vollständig überarbeitete und erweiterte Aufl. – Göttingen: Hogrefe, 2004

Hartmann, H.A.; Haubl, Rolf; Neuberger, Oswald; et al. (1984): Diagnostische Probleme psychologischer Begutachtung. In: Hartmann, H.A.; Haubl, R. (Hrsg.): Psychologische Begutachtung. Probleme und Praxisfelder. – München: Urban & Schwarzenberg, 1984

Hossiep, Rüdiger (1995): Berufseignungsdiagnostische Entscheidungen. Zur Bewährung eignungsdiagnostischer Ansätze. – Göttingen: Hogrefe, 1995

Hossiep, Rüdiger (1996): Psychologische Tests – Die vernachlässigte Dimension in Assessment Centern. In: Sarges, Werner (Hrsg.): Weiterentwicklungen der Assessment Center- Methode. – Göttingen: Verlag für Angewandte Psychologie, 1996

Hossiep, Rüdiger (2000): Konsequenzen aus neueren Erkenntnissen zur Potentialbeurteilung. In: Rosenstiel, Lutz; Lang-von Wins, Thomas (Hrsg.): Perspektiven der Potentialbeurteilung. – Göttingen: Verlag für Angewandte Psychologie, 2000

Hossiep, Rüdiger; Mühlhaus, Oliver (2005): Personalauswahl und –entwicklung mit Persönlichkeitstests. – Göttingen: Hogrefe (Praxis der Personalpsychologie-Band 6), 2005

Hossiep, Rüdiger; Paschen, Michael; Mühlhaus, Oliver (2000): Persönlichkeitstests im Personalmanagement. – Göttingen: Hogrefe, 2000

Ibrahimovic, Nedina; Bulheller, Stephan (2008): 5 DDP. Fünf Dimensionen der Persönlichkeit. Ein Verfahren zur Erfassung relevanter Eigenschaften der berufsbezogenen Persönlichkeit. Manual. – 1. Aufl. – München: Pearson, 2008

James, William (1981): The Principles of Psychology (Vol.1). – Cambridge: Havard University Press, 1981

Kanning, Uwe Peter; Pöttker, Jens; Klinge, Katharina (2008): Personalauswahl. Leitfaden für die Praxis. – Stuttgart: Schäffer-Poeschel, 2008

Kirby, Linda, K.; Myers, Katharine D. (2001): Einführung in Typen. – 6. Aufl. – Mountain View, USA: CPP, 2001

Kleinbeck, Uwe (1991): Die Wirkung von Zielsetzungen auf die Leistung. In: Schuler, Heinz (Hrsg.): Beurteilung und Förderung beruflicher Leistung. – Göttingen: Verlag für Angewandte Psychologie, 1991

Krohne, Heinz Walter; Hock, Michael (2007): Konstruktion und Überprüfung von Testverfahren. In: Heuer, Herbert; Rösler, Frank; Tack, Werner H. (Hrsg.): Psychologische Diagnostik. Grundlagen und Anwendungsfelder. – Stuttgart: Kohlhammer, 2007

Lienert, Gustav; Raatz, Ulrich (1998): Testaufbau und Testanalyse. – 6.Aufl. – Weinheim: Psychologie Verlags Union, 1998

Nußbaum, Albert; Neumann, Bodo (1995): Jede Entwicklung geht vom Menschen aus – Human Resources Management als unternehmerische Aufgabe. In: Jochmann, Walter (Hrsg.): Personalberatung intern. Philosophien, Methoden und Resultate führender Beratungsunternehmen. – Göttingen: Verlag für Angewandte Psychologie, 1995

Obermann, Christoph (2009): Assessment Center. Entwicklung, Durchführung, Trends – Mit originalen AC-Übungen, 4., vollständig überarb. und erweiterte Aufl. – Wiesbaden: Gabler, 2009

Paulhus, Del L. (1986): Self-Deception and Impression Management in Test Responses. In: Angleitner, Alois; Wiggins, Jerry S. (Hrsg.): Personality Assessment via Questionnaire. – Berlin: Springer, 1986

Rentzsch, Kathrin; Schütz, Astrid (2009): Psychologische Diagnostik. Grundlagen und Anwendungsperspektiven. – Stuttgart: Kohlhammer Urban Taschenbücher (Band 6), 2009

Schelten, Andreas (1997): Testbeurteilung und Testerstellung. Grundlagen der Teststatistik und Testtheorie für Pädagogen und Ausbilder in der Praxis. – 2., durchges. und erw. Aufl. – Stuttgart: Steiner, 1997

Schmidt, K.H. (1980): Zur rechtlichen Zulässigkeit verschiedener Formen des Personalinterviews. Der Betrieb (50). – 1980

Schuler, Heinz (2000): Psychologische Personalauswahl. Einführung in die Berufseignungsdiagnostik. – 3., unveränderte Aufl. – Göttingen: Hogrefe, 2000

Schuler, Heinz; Funke, Uwe (1995): Diagnose beruflicher Eignung und Leistung. In: Schuler, Heinz (Hrsg.): Lehrbuch Organisationspsychologie. – 2.Aufl. – Bern: Huber, 1995

Schumann, Siegfried (2005): Persönlichkeit. Eine vergessene Größe der empirischen Sozialforschung. – 1.Aufl. – Wiesbaden: Verlag für Sozialwissenschaftler, 2005

Skibba, Katrin (2006): Personalauswahl gemäß DIN 33430. Nutzenpotentiale für Unternehmen. – Saarbrücken: VDM, 2006

Wagner, Dieter (1995): Eignungsdiagnostik im Managementbereich. Personalpolitische Ausgangslage. In: Sarges, Werner (Hrsg.): Management-Diagnostik. – 2., vollständig überarb. und erweiterte Aufl. – Göttingen: Hogrefe, 1995

Weuster, Arnulf (2004): Personalauswahl. Anforderungsprofil, Bewerbersuche, Vorauswahl und Vorstellungsgespräche. – Wiesbaden: Gabler, 2004

Wiggins, Jerry S. (1973): Personality and Prediction: Principles of Personality Assessment – Malabar, Florida: Addison- Wesley, 1973

Wottawa, Heinrich (1980): Grundriß der Testtheorie. – München: Juventa, 1980

Wottawa, Heinrich; Hossiep, Rüdiger (1987): Grundlagen psychologischer Diagnostik. Eine Einführung. – Göttingen: Hogrefe, 1987

Ziegler, Beate (2002): NEO- Fünf- Faktoren- Inventar (NEO-FFI). In: Kanning, Uwe Peter; Holling, Heinz (Hrsg.): Handbuch personaldiagnostischer Instrumente. Göttingen: Hogrefe, 2002

Internetquellen

HIPA: Wissenswertes für Interessierte: Psychologie

URL: <http://www.hipa.at/psycho/eigenschaftstheorien.htm>, verfügbar am 12.06.2010

Erwin Ebermann, Institut für Kultur- und Sozialanthropologie:

URL: <http://www.univie.ac.at/ksa/elearning/cp/quantitative/quantitative-59.html>, verfügbar am 03.08.2010

Erwin Ebermann, Institut für Kultur- und Sozialanthropologie:

URL: <http://www.univie.ac.at/ksa/elearning/cp/quantitative/quantitative-96.html>, verfügbar am 03.08.2010

Erwin Ebermann, Institut für Kultur- und Sozialanthropologie:

URL: <http://www.univie.ac.at/ksa/elearning/cp/quantitative/quantitative-108.html>, verfügbar am 04.08.2010

Metaberatung GmbH: International Business Psychology

URL: http://www.metaberatung.de/05_Tests_und_Methoden/HDS.htm, verfügbar am 03.08.2010

Prof. Dr. Martin Kersting:

URL: <http://www.kersting-internet.de/start-p.html>, verfügbar am 19.08.2010

TU Wien: Tabellen von Verteilungen: Quantile, kritische Werte

URL: http://www.statistik.tuwien.ac.at/public/dutt/vorles/mb_wi_vt/node98.html, verfügbar am 23.07.2010

Zeitschriften/ Magazine/ Broschüren

Francis, Leslie J.; Jones, Susan H. (2000): The Relationship Between the Myers-Briggs-Type Indicator and the Eysenck Personality Questionnaire Among Adult Churchgoers. In: Pastoral Psychology – 2000 Vol. 48, No. 5, S. 377 - 386

Fruhner,R.; Schuler, Heinz; Funke, U.; et al. (1991): Einige Determinanten der Bewertung von Personalauswahlverfahren. In: Zeitschrift für Arbeits- und Organisationspsychologie. – Göttingen: Hogrefe. – 1991 Vol. 35, No. 4, S. 170 - 178

Furnham, Adrian; Crump, John (2005): Personality Traits, Types, and Disorders: An Examination oft the Relationship Between Three Self-Report Measures. In: European Journal of Personality. – London: John Wiley & Sons. – 2005 No. 19, S. 167 – 184

Hampel, R.; Klinkhammer,F. (1978): Verfälschungstendenzen beim Freiburger Persönlichkeitsinventar in einer Bewerbungssituation. In: Psychologie und Praxis. – Göttingen: Hogrefe. – 1978 No. 22, S. 58 – 69

Hogan, Robert; Hogan, Joyce; Roberts, Brent W. (1996): Personality Measurement and Employment Decisions. In: American Psychologist. – 1996: Vol. 51, No. 5, S. 469 – 477

McCrae, Robert R.; Costa, Paul T. (1989): Reinterpreting the Myers-Briggs Type Indicator from the perspective of the Five-Factor Model of Personality. In: Journal of Personality. – 1989 No. 57, S. 17 – 40

O'Meara, Daniel P. (1994): Personality tests raise questions of legality and effectiveness. In: Human Resource Magazine. – London: Haymarket Media. – 1994 No.39, S. S.97 – 100

Ruch, Willibald (1999): Die revidierte Fassung des Eysenck Personality Questionnaire und die Konstruktion des deutschen EPQ-R bzw. EPQ-RK. In: Zeitschrift für Differentielle und Diagnostische Psychologie. – Bern: Hans Huber. – 1999 Vol. 20, No. 1, S. 1 – 24

Schmidt, Frank L.; Hunter, John E. (1998): The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. In: Psychological Bulletin. – Washington: American Association – 1998 Vol. 124, No. 2, S. 262 – 274

Sonstiges

Furnham, Adrian; Moutafi, Joanna; Crump, John (2003): The relationship between the Revised NEO-Personality Inventory and the Myers-Briggs Type Indicator. Social Behavior and Personality. – Publikation, 2003

URL: http://findarticles.com/p/articles/mi_qa3852/is_200301/ai_n9213480/, verfügbar am 20.05.2010

Hänsgen, Klaus-D. (1999): Stand und einige Perspektiven der Psychodiagnostik. Forschungsbericht Nr. 147. – Freiburg: Psychologisches Institut der Universität Freiburg, 1999

Hossiep, Rüdiger (2007): Eignungsdiagnostik: Basiswissen und dessen praktische Umsetzung. - Mit Diagnostik zum Erfolg; Wirtschaftspsychologischer Intensivtag. – Schloss Montabaur – Sektion Wirtschaftspsychologie im Berufsverband Deutscher Psychologinnen und Psychologen e. V. (BDP) in Kooperation mit der Landesgruppe Nordrhein-Westfalen im BDP e. V.

Klein, Frank J. (1982): Die Rechtmäßigkeit psychologischer Tests im Personalbereich. (Diss. iur. Univ. Münster) – Gelsenkirchen: Mannhold, 1982

Eigenständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt zu haben.

Halsbrücke, den 24.08.2010

Katrin Hillgruber